

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 02:35:21 ; Search time 2357.99 Seconds
(without alignments)
19464.949 Million cell updates/sec

Title: US-09-820-788A-1
Perfect score: 1537
Sequence: 1 cctgctggtctctgtgcc.....aaaaaaaaaaaaaaaaaaaaa 1537

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hct:*
9: gb_est1:*
10: gb_est2:*
11: gb_hct:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gas_hum:*
18: em_gas_inv:*
19: em_gas_pin:*
20: em_gas_vrt:*
21: em_gas_fun:*
22: em_gas_nam:*
23: em_gas_mus:*
24: em_gas_pro:*
25: em_gas_rod:*
26: em_gas_pig:*
27: em_gas_vrl:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	950.6	61.8	1201	13	BX422591 BX422591
C 2	744.2	48.4	1071	13	BX432877 BX432877
C 3	731.2	47.6	767	14	CK032844 AGENCOURT
C 4	721.6	46.9	803	14	CK032848 AGENCOURT

C 5	721.2	46.9	798	14	CF271947
C 6	720.8	46.9	780	14	CK032843
C 7	719.6	46.8	814	14	CK032845
C 8	714.8	46.5	804	14	CK032846
C 9	709	46.1	807	14	CK032847
C 10	696.6	45.3	2429	11	AK090296
C 11	695	45.2	796	14	CK130223
C 12	688.2	44.8	791	14	CF271949
C 13	677.4	44.1	1646	11	AK004915
C 14	663.4	43.2	1201	13	BX422592
C 15	652.8	42.5	1524	11	AK078880
C 16	651.8	42.4	774	14	CK032849
C 17	609.8	39.7	982	12	BG743095
C 18	583	37.9	583	13	BX096000
C 19	583	37.9	586	14	CB162434
C 20	576	37.5	617	12	BM682376
C 21	575.4	37.4	1088	13	BM682376
C 22	564.2	36.7	985	13	BM682376
C 23	563	36.6	587	12	BM717845
C 24	559	36.4	3079	11	AK004933
C 25	529.8	34.5	880	12	BI217581
C 26	526.4	34.2	809	9	AI987772
C 27	525.4	34.2	796	9	AI255822
C 28	524.8	34.1	902	13	BI328628
C 29	509.4	33.1	808	12	BI328628
C 30	508.4	33.1	631	14	CB456776
C 31	508	33.1	798	12	BI247704
C 32	504	32.8	1066	14	CB195223
C 33	503.6	32.8	803	9	AI132450
C 34	500.4	32.6	736	12	BI103579
C 35	499.6	32.5	793	9	AI256232
C 36	495.6	32.2	788	9	AI266991
C 37	494.2	32.2	798	9	AI194824
C 38	494	32.1	782	9	AI1326364
C 39	493	32.1	806	9	AI116003
C 40	492.6	32.0	787	9	AI314456
C 41	491.6	32.0	741	12	BI331279
C 42	491.6	32.0	976	13	BI331279
C 43	488.8	31.8	797	12	BI217032
C 44	488	31.8	734	9	AI046427
C 45	487.2	31.7	641	10	BF534348

ALIGNMENTS

RESULT 1
BX422591/c
LOCUS
DEFINITION BX422591 Homo sapiens FETAL LIVER Homo sapiens CDNA clone
CSODM008Y007 3-PRIME, mRNA sequence.
ACCESSION BX422591.1 GI:30659345
VERSION BX422591
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
TITLE Li.W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10301.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODM008AH04NP1cluster=10301.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DM008AH04NP1.

FEATURES

source

```
1. 1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM008Y007"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
```

ORIGIN

```
Query Match 61.8%; Score 950.6; DB 13; Length 1201;
Best Local Similarity 98.5%; Pred. No. 5.3e-113;
Matches 980; Conservative 9; Mismatches 3; Indels 3; Gaps 3;

Qy 435 CCCTTTCCGCCCAACGGTCTCTTGACAAAGCCGTTGAGCAACGTTGATCGCTTCCCTCACC 494
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
998 CCCTTTCCGCCCAACGGTCT-TTGAMAAASCCGTGAGCAACGTTGATSSCTTCCCTCMAC 940
Qy 495 TGGGGCGCGCTTCGACTAGCAGACCTCGCTTCTCAGGCTGCTGGACTAGCTCAG 554
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
939 TGGGGCGCGCTTCGACTAGCAGACCTCGCTTCTCAGGCTGCTGGACTAGCTCAG 880
Qy 555 GAGGACTGAAGGAGGAGTCCGGCTTTCTGGCGGAGGTGCTGAATGCTGCCCGTCTC 614
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
879 GAGGACTGAAGGAGGAGTCCGGCTTTCTGGCGGAGGTGCTGAATGCTGCCCGTCTC 820
Qy 615 CTGCATAT-CCGAGCGCTGGTGGCAAGGTCTTACGCTTCC-AAAGGCTTTCCTGACCC 672
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
819 CTGCATATCCCGAGCGCTGGTGGCAAGGTCTTACGCTTCCMAAAGGCTTTCCTGACCC 760
Qy 673 AGCTGGATGAGTCTAACTGAGCAGAGTACCTGGAGCCAGCCAGCCAGCCCGCCGAG 732
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
759 AGCTGGATGAGTCTAACTGAGCAGAGTACCTGGAGCCAGCCAGCCCGCCCGAG 700
Qy 733 ACCTGACTGAGGCTTCTCTGGCAGAGATGGAGAGGCCAAGGGAAACCTTGAGAGCAGCT 792
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
699 ACCTGACTGAGGCTTCTCTGGCAGAGATGGAGAGGCCAAGGGAAACCTTGAGAGCAGCT 640
Qy 793 TCAATGATGAGAACCTCGCCATAGTGGTGGTGAACCTTCTCTGCGGAGATGGTGAACA 852
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
639 TCAATGATGAGAACCTCGCCATAGTGGTGGTGAACCTTCTCTGCGGAGATGGTGAACA 580
Qy 853 CTGAGACCACTGGCTGGGCTTCTCTGCTCATGATCCTATCATCCGATGTGAGCGCC 912
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
579 CTGAGACCACTGGCTGGGCTTCTCTGCTCATGATCCTATCATCCGATGTGAGCGCC 520
Qy 913 GTGTCCAAAGAGATCGACGACGTGATAGGCGAGTGGCGGACGACAGATGGTGAACC 972
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
519 GTGTCCAAAGAGATCGACGACGTGATAGGCGAGTGGCGGACGACAGATGGTGAACC 460
Qy 973 AGGCTCACATGCCCTTACCACTGCGGTGATTATGAGTGCAGCGCTTTCGGGACATCG 1032
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
459 AGGCTCACATGCCCTTACCACTGCGGTGATTATGAGTGCAGCGCTTTCGGGACATCG 400
Qy 1033 TCCCTCTGGTGTGACCCATATGATCCCGTGATGATCGAGTACAGGGCTTCGCGATCC 1092
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
399 TCCCTCTGGTGTGACCCATATGATCCCGTGATGATCGAGTACAGGGCTTCGCGATCC 340
Qy 1093 CTAAGGGAACGACACTCATATCAACCACTGTCTATCGGTCTGAGGATGAGCGCTCTGGG 1152
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
339 CTAAGGGAACGACACTCATATCAACCACTGTCTATCGGTCTGAGGATGAGCGCTCTGGG 280
Qy 1153 AGAAGCCCTTCGGCTTCCACCCCGGAACAATTCTTGATGCCAGGCGCACTTTGTGAAGC 1212
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
279 AGAAGCCCTTCGGCTTCCACCCCGGAACAATTCTTGATGCCAGGCGCACTTTGTGAAGC 220
```

```
Qy 1213 CGGAGGCTTCTCTGCTTCTTCTCAGCAGCGCGCTGATGCTCGGGAGAGCCCTGGGCC 1272
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
219 CGGAGGCTTCTCTGCTTCTTCTCAGCAGCGCGCTGATGCTCGGGAGAGCCCTGGGCC 160
Qy 1273 GCATGAGGCTTCTCTGCTTCTTCTCAGCTCCCTGCTGAGCACTTTCAGCTTCTCGGTGCCCA 1332
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
159 GCATGAGGCTTCTCTGCTTCTTCTCAGCTCCCTGCTGAGCACTTTCAGCTTCTCGGTGCCCA 100
Qy 1333 CTGGACAGCCCGCGCCAGCCACCATGGTCTTCTTCTGCTTCTTCTGACCCCATCCCT 1392
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
99 CTGACAGCCCGCGCCAGCCACCATGGTCTTCTTCTGCTTCTTCTGACCCCATCCCT 40
Qy 1393 ATGAGCTTGTGCTGCTGCCCGCTAGAAATGGGTA 1427
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
39 ATGAGCTTGTGCTGCTGCCCGCTAGAAATSCGTA 5
```

```
RESULT 2
BX432877/c
LOCUS BX432877 1071 bp mRNA linear EST 15-MAY-2003
DEFINITION BX432877 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CS0DM004YJ13 3-PRIME, mRNA sequence.
ACCESSION BX432877
VERSION BX432877.1 GI:30777112
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1071)
AUTHORS Li,W.B., Gruber,C., Jesses,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10301.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAH018ZE10NP1&cluster=10301.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue genoscope sequence ID : CS0BAH018ZE10NP1.
FEATURES
Location/Qualifiers
1. 1071
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM004YJ13"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
```

```
Query Match 48.4%; Score 744.2; DB 13; Length 1071;
Best Local Similarity 90.9%; Pred. No. 1.8e-86;
Matches 798; Conservative 38; Mismatches 37; Indels 5; Gaps 5;

Qy 563 GAAGAGAGTTCGGGCTTCTGCGGAGGTGCTGAATGCTGCCCGCTCTCTCTGCAAT 622
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
911 GACTGAAGAGAGTTCGGGCTTCTGCGGAGGTGCTGAATGCTGCCCGCTCTCTCTGCAAT 852
Qy 623 CCCAGCGCTGGCTGGCAAGGTCTTACGCTTCCAAAAGGCTTCTTGACCCAGCTGGATGA 682
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
851 CCCAGCGCTGGCTGGCAAGGTCTTACGCTTCCAAAAGGCTTCTTGACCCAGCTGGATGA 793
```


KEYWORDS	EST.	947	GCTGGCGGACCCAGAGATGGGTGACAGGCTCACATGCGCTACACACCTGCGGTGATTCA	1006
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens	505	GCTGGCGGACCCAGAGATGGGTGACAGGCTCACATGCGCTACACACCTGCGGTGATTCA	446
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 798)	1007	TCAGCTGCAGCGCTTTGGGAGACATGTCCTCCCTGGGTGTGACCCATATGACATCCCGTGA	1066
TITLE	NIH-MGC http://mgs.nci.nih.gov/.			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)	445	TGAGGTGCAGCGCTTTGGGAGACATGTCCTCCCTGGGTGTGACCCATATGACATCCCGTGA	386
COMMENT	Unpublished (1999)			
	On Aug 12, 2003 this sequence version replaced gi:33627859.	1067	CATCGAAGTACAGGCTTCCCGCATCCCTAAGGGAACGACACTCATCACCAACCTGTGCATC	1126
	Contact: Daniela S. Gerhard, Ph.D.			
	National Cancer Institute / NIH	385	CATCGAAGTACAGGCTTCCCGCATCCCTAAGGGAACGACACTCATCACCAACCTGTGCATC	326
	Bldg. 31 Rm10A07 Bethesda, MD 20892			
	Email: cgabbs-remail.nih.gov	1127	GCTGCTGAAGGATGAGGCGCTCTGGGAGAGCCCTTCGCTCCACCCGACACATTCCT	1186
	Tissue Procurement: Narayan Bhat			
	cDNA Library Preparation: Bhat Laboratory	325	GCTGCTGAAGGATGAGGCGCTCTGGGAGAGCCCTTCGCTCCACCCGACACATTCCT	266
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Agencourt Bioscience Corporation	1187	GCATGCCAGGCGCACTTTGTGAAGCCGAGAGCCCTTCCTGCCTTCTCAGCAGGCGCGCG	1246
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:	265	GGATGCCAGGCGCACTTTGTGAAGCCGAGAGCCCTTCCTGCCTTCTCAGCAGGCGCGCG	206
	http://image.llnl.gov			
	Plate: IRBK4 row: c column: 11	1247	TGCATGCTCTGGGAGAGCCCTTCGCGCATGGAGCTCTTCCTCTTCTTCCCTCCCTGCT	1306
	High quality sequence start: 8			
	High quality sequence stop: 748.	205	TGCATGCTCTGGGAGAGCCCTTCGCGCATGGAGCTCTTCCTCTTCTTCCCTCCCTGCT	146
FEATURES	Location/Qualifiers			
source	1..798			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:7002157"			
	/tissue="mixed"			
	/lab_host="DH5A (T1 phage-resistant)"			
	/clone_lib="NIH MGC 195"			
	/note="Vector: pDNR-Dual; Site 1: loxp-Sall; Site 2:			
	loxP-HindIII; Clones from this library have been			
	PCR-amplified using gene-specific primers to contain the			
	complete open reading frame (based on known gene sequences			
	available from NCBI's RefSeq). Template for PCR is cDNA			
	derived from either pooled cytoplasmic polyA RNA from 30			
	cells lines or pooled total RNA from 10 different tissues			
	(from BD Biosciences/Clontech and Washington University).			
	PCR products are directionally cloned into the loxp sites			
	of the pDNR-Dual vector. Library constructed by Dr.			
	Narayan Bhat, Earl Bere III and Hongling Liao (Gene			
	Expression Laboratory, Research Technology Program, SAIC			
	Frederick, NCI-Frederick, Frederick, MD 21702). For			
	information on which gene each clone represents, please			
	visit our anonymous ftp site at			
	ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat			
	a Note: this is a NIH_MGC Library."			
ORIGIN				
Query Match	46.9%; Score 721.2; DB 14; Length 798;			
Best Local Similarity	98.1%; Pred. No. 1.8e-83;			
Matches 729; Conservative	0; Mismatches 14; Indels 0; Gaps 0;			
		707	CTGGACCCAGCCAGCCCGGAGACCTGACTGAGGCTTCTTGGGAGAGATGAGAA	766
		745	CNTGACCCAGCCAGCCCGGAGACCTGACTGAGGCTTCTTGGGAGAGATGAGAA	686
		767	GGCCAAAGGGGAACCTTGAGAGAGCTTCAATGATGAGAACCTGCGCATAGTGGTCTGA	826
		685	GGCCAAAGGGGAACCTTGAGAGAGCTTCAATGATGAGAACCTGCGCATAGTGGTCTGA	626
		827	CTGTCTCTGCGGAGATGAGACCACTGACAGCTGGGCTGGGCTCTCTCAT	886
		625	CCTGTCTCTGCGGAGATGAGACCACTGACAGCTGGGCTGGGCTCTCTCAT	566
		887	GATCTTACATCCGATGTCAGGCGCTGTCACAGGAGATCGAGCTATAGGCA	946
		565	GATCTTACATCCGATGTCAGGCGCTGTCACAGGAGATCGAGCTATAGGCA	506

745 CTTTCTGGCAGAGATGGAGAGCCCAAGGGGAACCTTGAGACAGCTTCAATGATGAGA 804
|
|
|
706 CTTTCTGGCAGAGATGGAGAGCCCAAGGGGAACCTTGAGACAGCTTCAATGATGAGA 647
|
|
|
805 ACCTGGCAGATAGTGGTGGCTGACCTGTTCTCTCCGGGATGGTGAACCTGACACCGC 864
|
|
|
646 ACCTGTGCATAGTGGTGGCTGACCTGTTCTCTCCGGGATGGTGAACCTGACACCGC 587
|
|
|
865 TGGCCTGGGGCTCTCTGCTCATGATCCTACATCCGATGTGACGCGCGGTGCCAACAGG 924
|
|
|
586 TGGCCTGGGGCTCTCTGCTCATGATCCTACATCCGATGTGACGCGCGGTGCCAACAGG 527
|
|
|
925 AGATGACGACGATGATAGGACAGGTGGCGACACGAGATGGTGGTGCACAGGCTCACATGC 984
|
|
|
526 AGATGACGACGATGATAGGACAGGTGGCGACACGAGATGGTGGTGCACAGGCTCACATGC 467
|
|
|
985 CTTACACCACTGCGGTGATTCATGAGTGCAGTGCAGCGCTTTGGGACATCGTCCCTGGTG 1044
|
|
|
466 CTTACACCACTGCGGTGATTCATGAGTGCAGTGCAGCGCTTTGGGACATCGTCCCTGGTG 407
|
|
|
1045 TGACCATATGACATCCCGTACATCGAAGTACAGGCTTCCGATCCCTAAGGGGAACGA 1104
|
|
|
406 TGACCATATGACATCCCGTACATCGAAGTACAGGCTTCCGATCCCTAAGGGGAACGA 347
|
|
|
1105 CACTCATCAACCACTGTCTATCGGTGCTGAAGATGAGCGCGTCTGGGAGAACCCCTTCC 1164
|
|
|
346 CACTCATCAACCACTGTCTATCGGTGCTGAAGATGAGCGCGTCTGGGAGAACCCCTTCC 287
|
|
|
1165 GCTTCCACCCGACCACTCTCGATGCCAGGCGCACTTTGTGAAGCGGAGCGCTTCC 1224
|
|
|
286 GCTTCCACCCGACCACTCTCGATGCCAGGCGCACTTTGTGAAGCGGAGCGCTTCC 227
|
|
|
1225 TGCTTTCTCAGCAGCGCGCGTGCATGCTCTCGGGAGCGCCCTGGCGCGCATGGAGCTCT 1284
|
|
|
226 TGCTTTCTCAGCAGCGCGCGTGCATGCTCTCGGGAGCGCCCTGGCGCGCATGGAGCTCT 167
|
|
|
1285 TCCTCTTTTCACTCTCGTGCAGCACTTCACTGCTTCTCGGTGCCACTTGGACAGCCCGC 1344
|
|
|
166 TCCTCTTTTCACTCTCGTGCAGCACTTCACTGCTTCTCGGTGCCACTTGGAGCGCCC 107
|
|
|
1345 GGCCGACGACCACTGTGCTTTGCTTCTGTTGACCCGATCCCTATGAGCTTTGTG 1404
|
|
|
106 GGCCGACGACCACTGTGCTTTGCTTCTGTTGACCCGATCCCTATGAGCTTTGTG 47
|
|
|
1405 CTGTGCGCGCTAGATGGGTACCTAGTCCCGCAGCTGTCTCC 1448
|
|
|
46 CTGTGCGCGCTAGATGGGTACCTAGTCCCGCAGCTGTCTCC 3

RESULT 8
CK032846/c
LOCUS
DEFINITION
IMAGE:7002155 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
CK032846 804 bp mRNA linear EST 26-NOV-2003
AGENCOURT_15196952 NIH_MGC_195 Homo sapiens cdna clone
IMAGE:7002155 5', mRNA sequence.
CK032846
CK032846.1 GI:38558770
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 804)
NIH-MGC <http://imgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Narayan Bhat
cdna Library Preparation: Bhat Laboratory

cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: IRBK4 row: c column: 09
High quality sequence start: 12
High quality sequence stop: 686.

FEATURES

Location/Qualifiers
1..804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7002155"
/tissue_type="mixed"
/lab_host="DH5A (TI phage-resistant)"
/clone_lib="NIH_MGC_195"
/note="Vector: pDNR-Dual; Site 1: loxp-Sall; Site 2:
loxP-HindIII; Clones from this library have been
PCR-amplified using gene-specific primers to contain the
complete open reading frame (based on known gene sequences
available from NCBI's RefSeq). Template for PCR is cDNA
derived from either pooled cytoplasmic polyA RNA from 30
cells lines or pooled total RNA from 10 different tissues
(from BD Biosciences/Clontech and Washington University).
PCR products are directionally cloned into the loxp sites
of the pDNR-Dual vector. Library constructed by Dr.
Narayan Bhat, Earl Bere III and Hongling Liao (Gene
Expression Laboratory, Research Technology Program, SAIC
Frederick, NCI-Frederick, Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK_presv.dat
a Note: this is a NIH_MGC Library."

ORIGIN

Query Match 46.5%; Score 714.8; DB 14; Length 804;
Best Local Similarity 98.0%; Pred. No. 1.2e-82;
Matches 744; Conservative 0; Mismatches 13; Indels 2; Gaps 2;
QY 690 ACTGAGCAGGATGACCTGGGACCCAGCCAGCCCGGAGACCTGACTGAGGCGCTTC 749
|
|
|
DB 760 ACTGAGCAGGATGACCTGGGACCCAGCCAGCCCGGAGACCTGACTGAGGCGCTTC 703
|
|
|
QY 750 CTGGCAGAGATGAGAGCCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACCTG 809
|
|
|
DB 702 CTGGCAGAGATGAGAGCCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACCTG 643
|
|
|
QY 810 CGCATAGTGGTGGCTGACCTGTTCTCTGCGGGATGGTGACCACTCGACCGCTGGCC 869
|
|
|
DB 642 TGCATAGTGGTGGCTGACCTGTTCTCTGCGGGATGGTGACCACTCGACCGCTGGCC 583
|
|
|
QY 870 TGGGGCTCTCTCTCATGATCCTACATCCGATGTGACGCGCGGTGTCCACAGGAGATC 929
|
|
|
DB 582 TGGGGCTCTCTCTCATGATCCTACATCCGATGTGACGCGCGGTGTCCACAGGAGATC 523
|
|
|
QY 930 GACGACGTGATAGGCGAGGTGGCGACCCAGAGATGGGTGACCACTCACATGCCCTAC 989
|
|
|
DB 522 GACGACGTGATAGGCGAGGTGGCGACCCAGAGATGGGTGACCACTCACATGCCCTAC 463
|
|
|
QY 990 ACCACTGCGGTGATTTCATGAGGTGACGCGCTTTGGGGACATCGTCCCCCTGGGTGTGACC 1049
|
|
|
DB 462 ACCACTGCGGTGATTTCATGAGGTGACGCGCTTTGGGGACATCGTCCCCCTGGGTGTGACC 403
|
|
|
QY 1050 CATATGACATCCCGTGACATCGAAGTACAGGCGCTTCGGATCCCTTAAGGAGACGACATC 1109
|
|
|
DB 402 CATATGACATCCCGTGACATCGAAGTACAGGCGCTTCGGATCCCTTAAGGAGACGACATC 343
|
|
|
QY 1110 ATCAACCAACCTGCTCATCGTGTGAAGGATGAGGCGCTCTGGGAGAGCCCTTCGCGCTTC 1169
|
|
|
DB 342 ATCAACCAACCTGCTCATCGTGTGAAGGATGAGGCGCTCTGGGAGAGCCCTTCGCGCTTC 283
|
|
|
QY 1170 CACCCCGGAACACTTCTCTGGATGCCCGAGGCGCACTTTGTGAAGCGGAGGCGCTTCTCGCCT 1229

|||||
282 CACCCGACACTTCTCGATGCCAGGCGCACATTGTGAAGCCGAGGCGCTTCTCGCT 223
|||||
1230 TTCTCAGCAGCGCGCGTGCATGCTCGGAGAGCCCTCGGCGCGATGAGCTTTCTC 1289
|||||
222 TTCTCAGCAGCGCGCGTGCATGCTCGGAGAGCCCTCGGCGCGATGAGCTTTCTC 163
|||||
1250 TTCTTCACTCCCTGCTCAGACATTCAGCTTCTCGGCGCCACTGGACAGCCCGCGCC 1349
|||||
162 TTCTTCACTCCCTGCTCAGACATTCAGCTTCTCGGCGCCACTGGACAGCCCGCGCC 103
|||||
1350 AGCCACCATGCTGCTTTCTGCTTCTGCTGACCCCATCCCTATGAGCTTTGCTGTG 1409
|||||
102 AGCCACCATGCTGCTTTCTGCTTCTGCTGACCCCATCCCTATGAGCTTTGCTGTG 43
|||||
1410 CCCCCTAGATGGGCTACCTAGTCCCGAGCTGCTCCC 1448
|||||
42 CCCCCTAGATGGGCTACCTGCTCCCAAGCTTCAGGC 4

RESULT 9
CK032847/c
LOCUS
DEFINITION AGENCOURT 15196920 NIH MGC 195 Homo sapiens cDNA clone
IMAGE:7002154 5', mRNA sequence.
ACCESSION CK032847
VERSION CK032847.1 GI:38558771
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 807)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabp-remail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: Bhat Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: IRBK4 row: c column: 08
High quality sequence start: 9
High quality sequence stop: 665.
Location/Qualifiers
1. 807
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7002154"
/tissue_type="mixed"
/lab_host="DH5A (TI phage-resistant)"
/clone_lib="NIH MGC 195"
/note="Vector: pDNR-Dual; Site 1: loxp-SalI; Site 2:
loxP-HindIII; Clones from this library have been
PCR-amplified using gene-specific primers to contain the
complete open reading frame (based on known gene sequences
available from NCBI's RefSeq). Template for PCR is cDNA
derived from either pooled cytoplasmic polyA RNA from 30
cells lines or pooled total RNA from 10 different tissues
(from BD Biosciences/Clontech and Washington University).
PCR products are directionally cloned into the loxp sites
of the pDNR-Dual vector. Library constructed by Dr.
Narayan Bhat, Earl Bere III and Hongling Liao (Gene
Expression Laboratory, Research Technology Program, SAIC
Frederick, NCI-Frederick, Frederick, MD 21702). For

FEATURES
source

information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK.presv.dat
a Note: this is a NIH_MGC library."

ORIGIN

Query Match 46.1%; Score 709; DB 14; Length 807;
Best Local Similarity 99.2%; Pred. No. 6.5e-82;
Matches 723; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 710 GGACCCAGCCAGCCCGCCGAGACCTTCAATGATGAGAACCTGTGCATAGTGGTGCACCT 829
Db 744 GGGACCCGCGCCAGCCCGCCGAGACCTGA-TGAGGCGCTTCTGCGAGAGATGGAGAGGC 686
Qy 770 CAAGGGGAACCTCAGAGCAGCTTCAATGATGAGAACCTGTGCATAGTGGTGCACCT 829
Db 685 CAAGGGGAACCTCAGAGCAGCTTCAATGATGAGAACCTGTGCATAGTGGTGCACCT 626
Qy 830 GTTCTCTCGCGGATGTTGACCACTCGACCACTGGCGCTGGCGCTTCTGCTCATGAT 889
Db 625 GTTCTCTCGCGGATGTTGACCACTCGACCACTGGCGCTGGCGCTTCTGCTCATGAT 566
Qy 890 CCTACATCCGATGTGAGCGCGCGTGTCCAACAGAGATCGACACCTGATAGGCGAGGT 949
Db 565 CCTACATCCGATGTGAGCGCGCGTGTCCAACAGAGATCGACACCTGATAGGCGAGGT 506
Qy 950 GCGCGGACAGAGATGGTGTGACAGGCTCATGCTGCTTACACCTGCGCTGATTCATGA 1009
Db 505 GCGCGGACAGAGATGGTGTGACAGGCTCATGCTGCTTACACCTGCGCTGATTCATGA 446
Qy 1010 GGTGACGCGCTTTGGGACATCGTCCCTGGGTGTGACCCCATATGACATCCCTGACAT 1069
Db 445 GGTGACGCGCTTTGGGACATCGTCCCTGGGTGTGACCCCATATGACATCCCTGACAT 386
Qy 1070 CGAAGTACAGGCTTCCGCGATCCCTAAGGGGAACGACACTCATACCACTGTATCGGT 1129
Db 385 CGAAGTACAGGCTTCCGCGATCCCTAAGGGGAACGACACTCATACCACTGTATCGGT 326
Qy 1130 GCTGAAGATGAGCGCGTCTGGGAGAGCCCTTCCGCTTCCACCCGAGACTTCTTCTGGA 1189
Db 325 GCTGAAGATGAGCGCGTCTGGGAGAGCCCTTCCGCTTCCACCCGAGACTTCTTCTGGA 266
Qy 1190 TGCCACAGGCGCACCTTTGTGAAGCGGAGGCGCTTCTGCTTTCTCAGCAGCGCGCGTC 1249
Db 265 TGCCACAGGCGCACCTTTGTGAAGCGGAGGCGCTTCTGCTTTCTCAGCAGCGCGCGTC 206
Qy 1250 ATGCTCGGGAGCGCCCTGCGCGCATGGAGCTTTCTCTTTCTTCACTCCCTGCTGCA 1309
Db 205 ATGCTCGGGAGCGCCCTGCGCGCATGGAGCTTTCTCTTTCTTCACTCCCTGCTGCA 146
Qy 1310 GCATTTCAGCTTCTCGGTGCCACTGGACAGCCCGCGCGCCAGCCACCATGCTGTTTTC 1369
Db 145 GCATTTCAGCTTCTCGGTGCCACTGGACAGCCCGCGCGCCAGCCACCATGCTGTTTTC 86
Qy 1370 TTTCTGCTGACCCCATCCCTCATGAGCTTTGTGCTGTGCGCGCTAGAAATGGGTACC 1429
Db 85 TTTCTGCTGACCCCATCCCTCATGAGCTTTGTGCTGTGCGCGCTAGAAATGGGTACC 26

1430 TAGTCCCA 1438
25 TAGTCCCA 17

RESULT 10
AK090296
LOCUS
DEFINITION Mus musculus 21 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:G630039L19 product:cytochrome P450, 2d22,
full insert sequence.
ACCESSION AK090296
VERSION AK090296.1 GI:26105861
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)

AK090296 2429 bp mRNA linear HTC 20-SEP-2003
Mus musculus 21 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:G630039L19 product:cytochrome P450, 2d22,
full insert sequence.
AK090296
AK090296.1 GI:26105861
HTC; CAP trapper.
Mus musculus (house mouse)

ORGANISM Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
PUBMED 11076861
5
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12042159
6 (bases 1 to 2429)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://phantom.gsc.riken.go.jp/
FEATURES Location/Qualifiers
source 1..2429
organism="Mus musculus"

misc_feature 45.3%; Score 696.6; DB 11; Length 2429;
Best Local Similarity 79.6%; Pred. No. 1.8e-80;
Matches 849; Conservative 0; Mismatches 214; Indels 4; Gaps 2;
QY 429 GGACGCCCTTTGCGCCCAACCGTCTCTTGGACAAAGCCGTGAGCAACGTGATCGCTCC 488
DB 423 GGACACCCCTTTTTCAGCCCTAACACCCCTATTGGACAAAGCAGTGTGTAACTGATCGCGTCC 482
QY 489 CTCACCTCGCGCGCGCTTCGAGTACGACGACCCCTCGCTTCTCAGGCTGCTGGAACCTA 548
DB 483 CTCCTCTATGCTCGCGCTTCGAGTACGACGACCCCTTCATCAGGCTACTGCGGCTTG 542
QY 549 GCTCAGGAGGGAGTAAAGGAGGAGTGGGCTTTCTGCGCGAGGTGCTGAATGCTGCCCC 608
DB 543 TTGAAGGAAACTTTTAAAGGAGGAGCTGGATTTCTTACCCATGTTCTGATGTTTCCCG 602
QY 609 GTCTCTCTGCATATCCAGCGCTGGCTGGCAAGGTCTTACGCTTCCAAAGGCTTTTCCTG 668
DB 603 ATGCTCTCTGCATATCCAGCGCTGGCTGGCAAGGTCTTACGCTTCCAAAGGCTTTTCCTG 662
QY 669 ACCCAGCTGGATGAGTCTTCTTCTGAGCAGCAGTACCTGCGGACCCACGCGGAGGCTTCC 728
DB 663 ACCATGTTGGATGAGTCTTCTTCTGAGCAGCAGTACCTGCGGACCCACGCGGAGGCTTCC 720
QY 729 CGAGACCTGACTGAGGCTTCTTCTGCGCAGAGTGGAGAGCCCAAGGGAGACCTCGAGAGC 788
DB 721 CGAGATTTGACTGATGCTTCTTCTGCTGAGGTGGAGAGCCCAAGGGAGTCTCGAGAGC 780
QY 789 AGCTTCAATGATGAGAACCTGCGCATAGTGGTGGTGTGACCTGTTCTCTGCGGGATGGTG 848
DB 781 AGCTTCAATGATGAGAACCTGCGCATAGTGGTGGTGTGACCTGTTCTCTGCGGGATGGTG 840
QY 849 ACCACCTCGACCCAGCTGCGCTGGGCGCTTCTGCTCAGTACCTGATCGGATGTCGAG 908
DB 841 ACCACCTCGACCCAGCTGCGCTGGGCGCTTCTGCTCAGTACCTGATCGGATGTCGAG 900
QY 909 CGCGCTGTCACACAGAGATCGACGACGATGATAGGCGAGTGGCGGACAGAGATGGGT 968
DB 901 CGCGGAGTACACAGAGAAATCGATGATGATAGGCGAGTGCAGTGTCCAGAGATGGCA 960
QY 969 GACAGGCTACATCCCTTACACCACTGCGGTGATTCATGAGGTGCAGCGCTTTGGGAC 1028
DB 961 GACAGGCTGCATGCGCTTACACCACTGCTGTCATTCATGAGGTGCAGCGCTTTGGGAC 1020
QY 1029 ATCGCTGCGCTGGGTGTCAGCCATATGATCCCTGATCCCTGATCGAGTACAGGCTTCCCG 1088
DB 1021 ATTCCTGCTTGGTGTACCTTACAGGACTTCTCGTGACATTTGAATACAGGCTTCTTT 1080
QY 1089 ATCCCTTAAGGGAAACGACACTCATACCAACCTGTGTGCTGCTGAAGATGAGGCGCTC 1148
DB 1081 ATCCCTTAAGGGAAACGACACTCATACCAACCTGTGTGCTGCTGAAGATGAGGCTGCT 1140
QY 1149 TGGGAGAGCCCTTCCGCTTCCACCCCGAACACTTCTCTGAGTGCAGGCGCACTTTG 1208

Db 1141 TGGGAGAAAGCCCTCTGCTTCATCCTGAACACTTCTCGATGCCAGGGCCACTTTGTG 1200
Qy 1209 AAGCGGAGGCTTCCTGCTTCTCAGCAGCGCCCTGATGCTCGGGAGCCCTG 1268
Db 1201 AAGCTGAGGCTTCATGCCATTTCTCAGCAGCGCCAGATCATGCTGGGGAGCCCTG 1260
Qy 1269 GCGCGCATGAGCTTCTCTCTTCTTCACTCCCTGCTGACGACTTCTCAGCTTCTCGGTG 1328
Db 1261 GCGCGCATGAGCTTCTCTCTTCTTCACTCCCTGCTGACGCTTCTAGCATCTCAGTG 1320
Qy 1329 CCCACTGACAGCCCGCCAGCAGCACCATGCTGCTTCTGCTGACCCCATCC 1388
Db 1321 CCGCATGACAGCCCGCCAGCAGCATCATGCGCTCTTAGGGCTCTGACAAACCCATGC 1380
Qy 1389 CCCTATGAGCTTGTGCTGCTGCGCCGCTAGATGGGTACCTAGTCCCGACCTGCTCCC 1448
Db 1381 CCCTACCAGCTCTGTGCTTGTGCGCCGCTAAGAGGATGATGACATCTCCTCACTGTGCT 1440
Qy 1449 TAGCCAGAGGCTCTAATGTACAATAAAGCAATGTGGTAGTTCCAAA 1495
Db 1441 T-GTGGGGTCTTAATGTGCAATAAAGCAATTTACTCTAACTAAA 1485

RESULT 11

CK130223
LOCUS
DEFINITION AGNCOURT_15196974 NIH_MGC_195 Homo sapiens cDNA clone
IMAGE:7002153 5', mRNA sequence.

ACCESSION

CK130223

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 796)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapsb-r@mail.nih.gov

Tissue procurement: Narayan Bhat

cDNA Library Preparation: Bhat Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: IRBK4 row: c column: 07

High quality sequence start: 8

High quality sequence stop: 666.

Location/Qualifiers

1. 796

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:7002153"

/tissue_type="mixed"

/lab_host="DHSA (T1 phage-resistant)"

/clone_lib="NIH_MGC_195"

/note="Vector: pDNR-Dual; Site_1: loxP-Sali; Site_2:

loxP-HindIII; Clones from this library have been

PCR-amplified using gene-specific primers to contain the

complete open reading frame (based on known gene sequences

available from NCI's RefSeq). Template for PCR is cDNA

derived from either pooled cytoplasmic polyA RNA from 30

cells lines or pooled total RNA from 10 different tissues

(from BD Biosciences/Clontech and Washington University).

PCR products are directionally cloned into the loxP sites

of the pDNR-Dual vector. Library constructed by Dr.

Narayan Bhat, Earl Bere III and Hongling Liao (Gene
Expression Laboratory, Research Technology Program, SAIC
Frederick, NCI-Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK.presv.dat
a Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.2%; Score 695; DB 14; Length 796;
Best Local Similarity 98.3%; Pred. No. 4.1e-80;
Matches 701; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
Qy 68 TGAGCGAGGTATGGGGCTAGAGCACTGGTGGCCCTGCGCTGATAGTGGGCATCTTCT 127
Db 20 TCATGCACTATGGGGCTAGAGCACTGGTGGCCCTGCGCTGATAGTGGGCATCTTCT 79
Qy 128 GCTCCTGCTGACCTGATGCAACCGCGCCAAACGCTGGGCTGACGCTACTCACAGGCC 187
Db 80 GCTCCTGCTGACCTGATGCAACCGCGCCAAACGCTGGGCTGACGCTACTCACAGGCC 139
Qy 188 CTTGCCACTGCGGGCTGGGCACTGCTGCACTGGGCTGCACTTCCAGAACACCATCTG 247
Db 140 CTTGCCACTGCGGGCTGGGCACTGCTGCACTGGGCTGCACTTCCAGAACACCATCTG 199
Qy 248 CTTGCCACTGCGGGCTGGGCACTGCTGCACTGGGCTGCACTTCCAGAACACCATCTG 307
Db 200 CTTGCCACTGCGGGCTGGGCACTGCTGCACTGGGCTGCACTTCCAGAACACCATCTG 259
Qy 308 GGTGCTGCTGCTCAATGGGCTGGGCGCGCTGCGGAGCGCTGGTGAACCAACGCGAGGA 367
Db 260 GGTGCTGCTGCTCAATGGGCTGGGCGCGCTGCGGAGCGCTGGTGAACCAACGCGAGGA 319
Qy 368 CACGCGCAACCGCGCTGTCGCCATCACCCAGATCCTGGGTTTGGGCGCGCTTCCCA 427
Db 320 CACGCGCAACCGCGCTGTCGCCATCACCCAGATCCTGGGTTTGGGCGCGCTTCCCA 379
Qy 428 AGGAGCGCCCTTTCGCCCAACCGCTCTCTTGGACAAAGCGCTGAGCAACGCTGCTGCTC 487
Db 380 AGGAGCGCCCTTTCGCCCAACCGCTCTCTTGGACAAAGCGCTGAGCAACGCTGCTGCTC 439
Qy 488 CTTCACTGCGGGCGCGCTTTCGAGTACGACGACCTCGCTTCTCAGGGCTGCTGGACCT 547
Db 440 CTTCACTGCGGGCGCGCTTTCGAGTACGACGACCTCGCTTCTCAGGGCTGCTGGACCT 499
Qy 548 AGCTCAGAGGACTGAGAGGAGTGGGCTTTCGCGGAGCTGCTGGAATGCTGCTCC 607
Db 500 AGCTCAGAGGACTGAGAGGAGTGGGCTTTCGCGGAGCTGCTGGAATGCTGCTCC 559
Qy 608 GGTCTCTGCTCATATCCAGCGCTGGCTGGCAAGGTCTCTACGCTTCCAAAAGGCTTCT 667
Db 560 GGTCTCTGCTCATATCCAGCGCTGGCTGGCAAGGTCTCTACGCTTCCAAAAGGCTTCT 619
Qy 668 GACCCAGCTGGATGAGCTGCTAACTGAGCAGAGATGACCTGGGACCCAGGCCGCC 727
Db 620 GACCCAGCTGGATGAGCTGCTAACTGAGCAGAGATGACCTGGGACCCAGGCCGCC 679
Qy 728 CCGAGACTGATGAGGCTTCTTGGCAGAGATGGAGAGCCAGGGGAACC 780
Db 680 CCGAGACTGATGAGGCTTCTTGGCAGAGATGGAGAGCCAGGGGAACC 732

RESULT 12

CF271949

LOCUS

DEFINITION

AGNCOURT_15196970 NIH_MGC_195 Homo sapiens cDNA clone

IMAGE:7002156 5', mRNA sequence.

CF271949

ACCESSION

VERSION

CF271949.2 GI:38558844

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 791)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
On Aug 12, 2003 this sequence version replaced gi:33627861.
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-t@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: Bhat Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: IRBk4 row: C column: 10
High quality sequence start: 10
High quality sequence stop: 671.
Location/Qualifiers
1..791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7002156"
/tissue_type="mixed"
/lab_host="DH5A (T1 phage-resistant)"
/clone_lib="NIH_MGC_195"
/notes="Vector: pDNR-Dual; Site 1: loxp-Sall; Site 2:
loxP-HindIII; Clones from this library have been
PCR-amplified using gene-specific primers to contain the
complete open reading frame (based on known gene sequences
available from NCBI's RefSeq). Template for PCR is cDNA
derived from either pooled cytoplasmic polyA RNA from 30
cells lines or pooled total RNA from 10 different tissues
(from BD Biosciences/Clontech and Washington University).
PCR products are directionally cloned into the loxp sites
of the pDNR-Dual vector. Library constructed by Dr.
Narayan Bhat, Earl Bere III and Hongling Liao (Gene
Expression Laboratory, Research Technology Program, SAIC
Frederick, NCI-Frederick, Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK_presV.dat
a Note: this is a NIH_MGC Library."

ORIGIN
Query Match 44.8%; Score 688.2; DB 14; Length 791;
Best Local Similarity 96.1%; Pred. No. 3.1e-79;
Matches 714; Conservative 0; Mismatches 27; Indels 2; Gaps 1;

QY 64 GTAGTCAGGAGTATGGGGCTAGAGCACTGGTCCCTGCGGTGATAGTGGCCATCT 123
DB 20 GTTATCAGTCGATATGGGGCTAGAGCACTGGTCCCTGCGGTGATAGTGGCCATCT 79
QY 124 TCCTGCTCTCTGGTGGACCTGATGTCACCGGCGCAACGCTGGCTGCACTACTCACCAG 183
DB 80 TCCTGCTCTCTGGTGGACCTGATGTCACCGGCGCAACGCTGGCTGCACTACTCACCAG 139
QY 184 GCCCTCTGTCACCTGCGCGGCTGGGCAACCTGCTGATGTGGACTTCCAGAACACACCAT 243
DB 140 GCCCTCTGTCACCTGCGCGGCTGGGCAACCTGCTGATGTGGACTTCCAGAACACACCAT 199
QY 244 ACTGCTTCGACCAAGTTGGCGCGCGCTTCGGGGAAGCTGTTCAAGCTGCAAGCTGGCTGA 303
DB 200 ACTGCTTCGACCAAGTTGGCGCGCGCTTCGGGGAAGCTGTTCAAGCTGCAAGCTGGCTGA 259
QY 304 CGCCGCTGCTGCTCAATGGCTGCGCGGCTGGCGAGGCGCTGGTACCCAGGCG 363
DB 260 CGCCGCTGCTGCTCAATGGCTGCGCGGCTGGCGAGGCGCTGGTACCCAGGCG 319

FEATURES
source
1..791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7002156"
/tissue_type="mixed"
/lab_host="DH5A (T1 phage-resistant)"
/clone_lib="NIH_MGC_195"
/notes="Vector: pDNR-Dual; Site 1: loxp-Sall; Site 2:
loxP-HindIII; Clones from this library have been
PCR-amplified using gene-specific primers to contain the
complete open reading frame (based on known gene sequences
available from NCBI's RefSeq). Template for PCR is cDNA
derived from either pooled cytoplasmic polyA RNA from 30
cells lines or pooled total RNA from 10 different tissues
(from BD Biosciences/Clontech and Washington University).
PCR products are directionally cloned into the loxp sites
of the pDNR-Dual vector. Library constructed by Dr.
Narayan Bhat, Earl Bere III and Hongling Liao (Gene
Expression Laboratory, Research Technology Program, SAIC
Frederick, NCI-Frederick, Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK_presV.dat
a Note: this is a NIH_MGC Library."

RESULT 13
AK004915
LOCUS
DEFINITION
Mus musculus adult male liver cDNA, RIKEN full-length enriched
library, clone:1300006E06 product:CYTOCHROME P450 2D2 (EC
1.14.14.1) (CYP1D2) (P450-DB2) (P450-CMP2) (DBRISOXINE
4-HYDROXYLASE) homolog [Rattus norvegicus], full insert sequence.
ACCESSION
AK004915 GI:12836460
VERSION
AK004915.1 GI:12836460
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Saeki, N., Carninci, P.,
Komno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, Y., Ohara, E., Watahiki, M.,
Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913

```


Db 859 ACTGACACAGGATGACCTGGGACCCAGCCAGCCGCCCTGACTGAGGCCTTC 918
QY 750 CTGGCAGATGAGAGGAGCCAGGAGCCCTGAGCAGCTTCAATGATGAGACCTG 809
Db 919 CTGGCAGATGAGAGGAGCCAGGAGCCCTGAGCAGCTTCAATGATGAGAA-CTG 977
QY 810 CGCATAGTGGTGGCTGACCTGTTCTCTGCGGAGTGGTGAC 850
Db 978 CGCATAGTGGTGGCTGACCTGTTCTCTGCGGAGTGGTGAC 1018

RESULT 15

AK078880

LOCUS

DEFINITION

Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:9030605E09 product:CYTOCHROME P450, 2D9 homolog [Mus musculus], full insert sequence.

ACCESSION

AK078880

VERSION

AK078880.1

KEYWORDS

HTC; CAP trapper;

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1

AUTHORS

Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

TITLE

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE

20499374

PUBMED

11042159

REFERENCE

2

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.

TITLE

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL

Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE

20530913

PUBMED

11076861

REFERENCE

4

AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE

Functional annotation of a full-length mouse cDNA collection

JOURNAL

Nature 409, 685-690 (2001)

REFERENCE

5

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL

Nature 420, 563-573 (2002)

REFERENCE

6

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayamizu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koye, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES

source

1..1624

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM.DB:9030605E09"

/db_xref="MGI:2357571"

/db_xref="taxon:10090"

/clone="9030605E09"

/sex="male"

/tissue_type="colon"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

50..1564

/note="unnamed protein product; CYTOCHROME P450, 2D9 homolog [Mus musculus] (SPR|Q921V1, evidence: FASTY, 87.7%ID, 100%length, match=1512)

putative"

/codon_start=1

/protein_id="BAC37440.1"

/db_xref="GI:26347583"

/translation="MELTGTDLMSVAFTVFILLVDLMHRRQSWTSVPPGVPMPVLNGLQVLDLNPYSLYQNRQYDFVSLQMAKPVVNRKMKKEVLLTCGEDTA DRPVPFHEHLGFKPRSQMIFAPYGPWEQRFRFSLSRNFGKRLSEWIKKA GHLCDATQTQGVINPNTKKATCNVIALIFARRFEYDPLYLRMLKYLEWISLKE LSGIPLVINTPELILHPLADKFLQSQSFALVDNLATENTTTWDPAQSPNLTQ AFLAIEKQGNPSSNDENLRNVLDLFTAGLITSTLSALLMLILHPDQRRV QCEIDVIGQVRHPEDMAQHPYTNVIEHVQFQDVLPHLPRTSRLDIEVDQFLI PKGTILLPNSSVHMDDTWKELRPFHPHLDQAQGHVKEAFITFSAGRRSCLGEP LARMEFLFTCLLQRFSSVDPQOPQSDHRVFSIMVAPSPYOLCAVIREQGH"

1603..1608

/note="putative"

1624

/note="putative"

polyA_signal

polyA_site

ORIGIN

Query Match

Best Local Similarity

Matches 1041; Conservative

0; Mismatches 357; Indels 154; Gaps 2;

QY 93 CTGTGCCCCCTGGCGGTGATAGTGGCCATCTTCCTGCTCTGTGGTGGACCTGATCCACCGG 152

Db 74 CTGTGGTCTGTGGCCATATTTCACAGTCATCTTCATATTAATGCTGGACCTGATCCACCGG 133

QY 153 CGCCAACTGGGTGACCGCTACTACACAGGCCCTTCCACTGCTGCGGCGGCTGGCAAC 212

Db 134 CGCCAGAGCTGGACTTCTTGTGTACCCACAGGCCCTGTGCCATGCGCTGTCTGGTAAAC 193

QY 213 CTGTGATGATGGACTTCCAGAACACACCACTACTGCTTCGACCACTGTCGCGCGCGCTTC 272

Db 194 CTGTGAGGTAGACCTAAATATATATGCTTACAGCTTGTACAGCTTCAAAACCGCTAT 253

QY 273 GGGGACGTGTTTACGCTGACGCTGGCTGGACGCGCGGTGGTGTGCTCAATGGCGTGGCG 332

Db 254 GGTCACTGTTTAGCCTACAGATGGCTGGAGGCTATGGTGTGTGATCAACAGAAATGAAG 313

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 28, 2004, 01:41:15 ; Search time 3825.13 Seconds
(without alignments)
17415.928 Million cell updates/sec

Title: US-09-820-788A-1
Perfect score: 1537
Sequence: 1 cctgcgtgctctgtgcc.....aaaaaaaaaaaaaaaaaaaa 1537

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sta.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sta.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rod.*
- 36: em_hgt_mam.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgt_hum.*
- 40: em_hgt_mus.*
- 41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1238.2	80.6	1567	6	AX687029	Sequence
2	1238.2	80.6	1567	9	HSDB1	X08006 Homo sapien
3	1238.2	80.6	1567	9	HUMCYPDB1	M20403 Human cytoc
4	1229.2	80.0	1568	6	AR084372	Sequence
5	1225.6	79.7	1568	6	AX687030	Sequence
6	1225.2	79.7	1566	6	AR084365	Sequence
7	1216.2	79.1	1568	6	A20907	A20907 debrisoquin
8	1211.2	78.8	1565	6	AX687033	Sequence
9	1204.4	78.4	1566	6	AR084374	Sequence
10	1201	78.1	1545	6	E15820	CDNA encodi
11	1181.4	76.9	1585	9	AF301911	AF301911 Macaca fu
12	1173.2	76.3	1494	6	E10868	CDNA encodi
13	1173.2	76.3	1494	6	AR399336	Sequence
14	1171.6	76.2	1494	6	E10647	Human cDNA
15	1171.6	76.2	1494	6	E10867	CDNA encodi
16	1171.6	76.2	1494	6	E10869	CDNA encodi
17	1171.6	76.2	1494	6	AR399335	Sequence
18	1171.6	76.2	1494	6	AR399337	Sequence
19	1170	76.1	1494	6	E10870	CDNA encodi
20	1170	76.1	1494	6	AR399338	Sequence
21	1165	75.8	1571	6	AR084373	Sequence
22	1159.4	75.4	1569	6	AR084375	Sequence
23	1103.8	71.8	1593	6	AX195173	Sequence
24	1097.4	71.4	1615	9	D29822	Callithrix
25	1094.8	71.2	1494	9	AY082602	Callithrix
26	1072.8	69.8	1218	9	HSP450	Human mRNA
27	1063	69.2	1195	6	AR380887	Sequence
28	1063	69.2	1195	9	HMP450DB1	Human cytoc
29	1061.2	69.0	1494	9	MFU38218	Macaca faec
30	973.6	63.3	1501	6	AR084376	Sequence
31	831.4	54.1	1615	4	BTCP2DA	B.taurus CY
32	830.2	54.0	1652	4	SSVD325H	Sus scrofa
33	828.6	53.9	1449	6	AX411044	Sequence
34	828.6	53.9	1449	9	HSDB1A	Human mRNA
35	808.8	52.6	1538	4	BTCP2D	B.taurus CY
36	788.2	51.3	1503	4	AB008785	Oryctolag
37	775.8	50.5	1665	4	DOGCP2D	Canis fami
38	766.8	49.9	1631	4	AB004268	Canis fami
39	763	49.6	1270	6	AX411045	Sequence
40	763	49.6	1270	9	HSDB1B	Human mRNA
41	760.6	49.5	1741	10	BC016256	Mus muscu
42	759.4	49.4	1503	4	AB008784	Oryctolag
43	747.2	48.6	1609	10	AF221525	Mus muscu
44	741.2	48.2	1778	10	RNU48219	Rattus norv
45	732.2	47.6	1632	10	AB008425	Rattus no

ALIGNMENTS

RESULT 1	AX687029	AX687029	Sequence 3 from Patent	DNA	linear	PAT 31-MAR-2003
AX687029	LOCUS	Sequence 3 from Patent	1567 bp	DNA	linear	PAT 31-MAR-2003
DEFINITION	AX687029	Sequence 3 from Patent	1567 bp	DNA	linear	PAT 31-MAR-2003
ACCESSION	AX687029	Sequence 3 from Patent	1567 bp	DNA	linear	PAT 31-MAR-2003
VERSION	AX687029.1	GI:29409533				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						

FEATURES		Location/Qualifiers		
source	1..1567		541	
	/organism="Homo sapiens"		541	
	/mol_type="mRNA"		541	
	/db_xref="taxon:9606"		541	
	/chromosome="22"		541	
	/map="22q11.2-qter"		541	
	/cell_type="hepatocyte"		541	
	/tissue_type="liver"		541	
	/clone_lib="lambda gt11"		541	
	1..1494		541	
CDS	/codon_start=1		541	
	/product="cytochrome P450 db1"		541	
	/protein_id="CAA30807.1"		541	
	/db_xref="GI:30451"		541	
	/db_xref="GOA:P10635"		541	
	/translation="MGLEALVPLAVIVAFILLLVDLMHRRQWAAARYPPGPLPLPGLG NLLHVDQNTFYCFQQLRRRFGDVSFLQALWTPVVLNGLAAVREALVTHGEDTADRP PVPIITQILGFRSQGLARYGPWRRQRFVSVLRLNLGLKXSLQWVTEEAACL CAAFANHSRPPRPNGLLDKAVSNVIALSTCGRRFEYDDPRFLRLDLAQELKEESG FLRELVNAVPLVLLHLPALAGKVLRFQKAPLTOLDLLEHTEHRMTWPAOPRDLTAPFL AEMERAKGNPSSFNENLRIVVADLFSAGMVTSTTLLAWGLLLMLHPDVORRVQOE LDDVIGQVRRPMPGQAHMPYITAVIHEVQRFGDIVPLGWTMTSRDIEVQGFRLPKG TTLITNLGSLVLDKDEAVWEKPRFHPHEHFLDAQHFKPEAFLPFSAGRACLGEPLAR MSLFLFTSLQLQHFSSVPTGQPRPSHHGVFAFLVSPSPYELCAVPR"		541	
	1546..1550		541	
	polya_signal	/note="putative"		541
	polyA_site	/note="polyA site"		541
	ORIGIN	Query Match	80.6%; Score 1238.2; DB 9; Length 1567;	541
Best Local Similarity		89.7%; Pred. No. 1.4e-205;	541	
Matches 1406; Conservative 0; Mismatches 8; Indels 153; Gaps 1;			541	
Qy		78	ATGGGGCTAGAGCACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCTGCTCTCTGGTG	137
Db		1	ATGGGGCTAGAGCACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCTGCTCTCTGGTG	60
Qy		138	GACCTGATGACCGGCGCAACGCTGGGCTGCACGCTACTACACAGGCCCTGCCACTG	197
Db		61	GACCTGATGACCGGCGCGCAACGCTGGGCTGCACGCTACTCCACAGGCCCTGCCACTG	120
Qy		198	CCGGGCTGGGCAACCTGCTGCATGTGGACTTCCAGAAACACACCACTACTGCTTCGACCA	257
Db		121	CCGGGCTGGGCAACCTGCTGCATGTGGACTTCCAGAAACACACCACTACTGCTTCGACCA	180
Qy		258	TTGGCGCGCGCTTTCGGGGAACGTGTTGACGCTGCAGCTGGCTGGACCGCGGTGGTCGTG	317
Db	181	TTGGCGCGCGCTTTCGGGGAACGTGTTGACGCTGCAGCTGGCTGGACCGCGGTGGTCGTG	240	
Qy	318	CTCAATGGGCTGGCGCGCTGGCGAGGCGCTGGTGACCCAGCGGAGGACACCGCGCGAC	377	
Db	241	CTCAATGGGCTGGCGCGCTGGCGAGGCGCTGGTGACCCAGCGGAGGACACCGCGCGAC	300	
Qy	378	CGCCCGCTGTGCCCATACCCAGATCTCTGGGTGTTGGGCGCGCTTCCCAA	428	
Db	301	CGCCCGCTGTGCCCATACCCAGATCTCTGGGTGTTGGGCGCGCTTCCCAA	360	
Qy	429	-----	428	
Db	361	CTGGCGCGCTATGGGCGCGCTGGCGAGAGAGAGCGGCTTCTCGTGTCCACCTTGGCG	420	
Qy	429	-----	428	
Db	421	AACCTGGGCTGGGCAAGAAGTCGCTGGAGCAGTGGGTGACCGAGAGGCGCGCTT	480	
Qy	429	-----GGAGCGCCCTTTCGCCCCCAACGGTCTCTTGGACAAA	464	
Db	481	TGTGGCGCGCTTTCGCCAACACCTCCGGAGCGCCCTTTCGCCCAACGGTCTCTTGGACAAA	540	
Qy	465	GCCGTGAGCAACGTGATGCCTCTCTCACCTGCGGGCGCGCTTCGAGTACGACGACCT	524	

HUMCYPDB1
LOCUS HUMCYPDB1 1567 bp mRNA linear PRI 02-NOV-1994
DEFINITION Human cytochrome P450 db1 mRNA, complete cds.
ACCESSION M20403.1
VERSION M20403.1 GI:181349
KEYWORDS cytochrome P450; debrisoquine 4-hydroxylase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1567)
Gonzalez, F.J., Vilbois, P., Hardwick, J.P., McBride, O.W.,
Nebert, D.W., Gelboin, H.V. and Meyer, U.A.
Human debrisoquine 4-hydroxylase (P450IID1): cDNA and deduced amino
acid sequence and assignment of the CYP2D locus to chromosome 22
Genomics 2 (2), 174-179 (1988)
3410476
MEDLINE 88314109
PUBMED
COMMENT Original source text: Human hepatocyte, cDNA to mRNA.
Draft entry and printed copy of sequence for [1] kindly provided by
D.W.Nebert, 15-JUL-1988.
FEATURES
source
1..1567
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="22q11.2-qter"
1..1567
/gene="CYP2D"
1..1567
/gene="CYP2D"
/product="CY P450 db1 mRNA"
1..1494
/gene="CYP2D"
/note="debrisoquine 4-hydroxylase"
/codon_start=1
/protein_id="AAAS2153.1"
/db_xref="GI:181350"
/db_xref="GB:000-119-832"
/translation="MGLRALVPLAVIVAIFLLVDLMHRRORWAARYPPGLPLPLGLG
NLLHVDFTQYPCQFLRRFGDFVSLQALWTVVLLGLAAVREALVTHGEDADRL
PVPIQIILGFRPSQVFLARYGPAMRQRFVSVTLNLGLKKSLEQWTEEAACL
CAAPANHSRPFPRNGLDKAVSNVIALTCGRFFEYDDPRFLRLDLAQLKEESG
FLREVLNAVPLLIHPALAKVILRQKAFILQDELLEHBMWDPAGPRDLFEAL
AEMKAKNPSSFNENLRIVADLFSAGWVITSLIWLGLLMLHPVQRRVQOE
IDDDVTGQRRPEMGDQAHMPYTTAVIEVQRFQDIVPLGTMHTSRDIEVQGFRI PKG
TTLITNLSVLKDAEWKPFPHFPELQAGHVFPEAFLPFSAGRRACLGRPLAR
MELFLFTSLIQHFSFVPTGQPRSHHGVEAFVSPSPYELCAVPR"
Chromosome 22.
ORIGIN
Query Match 80.6%; Score 1238.2; DB 9; Length 1567;
Best Local Similarity 89.7%; Pred. No. 1.4e-205;
Matches 1406; Conservative 0; Mismatches 8; Indels 153; Gaps 1;
QY 78 ATGGGGCTAGAGCACTGTGCGCCCTGGCGGTAGTGGCCATCTTCTGCTCTGTG 137
DB 1 ATGGGGCTAGAGCACTGTGCGCCCTGGCGGTAGTGGCCATCTTCTGCTCTGTG 60
QY 138 GACCTGATGCACCGGCGCAACGCTGGGTGACGCTACTACACAGGCGCCCTGCACTG 197
DB 61 GACCTGATGCACCGGCGCAACGCTGGGTGACGCTACTACACAGGCGCCCTGCACTG 120
QY 198 CCGGGCTGGGCAACCTGCTGATGTGACTTCCAGAACACACCATCTGCTTCAGCCAG 257
DB 121 CCGGGCTGGGCAACCTGCTGATGTGACTTCCAGAACACACCATCTGCTTCAGCCAG 180
QY 258 TTGCGGCGCGCTTCGGGAGCGTGTTCAGCCTGACGCTGGCTGACGCCGCGTGTGCTG 317
DB 181 TTGCGGCGCGCTTCGGGAGCGTGTTCAGCCTGACGCTGGCTGACGCCGCGTGTGCTG 240
QY 318 CTCAATGGGCTGGCGGCGCTGCGGAGGCGTGTGATCCACCGGAGGACACACCGCGAC 377
DB 241 CTCAATGGGCTGGCGGCGCTGCGGAGGCGTGTGATCCACCGGAGGACACACCGCGAC 300

QY 378 CGCCGCGCTGTGCGCCATCACCCAGATCCTGGGTTTTTGGGCGCGCTTCCCAA----- 428
DB 301 CGCCGCGCTGTGCGCCATCACCCAGATCCTGGGTTTTTGGGCGCGCTTCCCAAAGGGGTGTTTC 360
QY 429 ----- 428
DB 361 CTGGCGGCTATGGGCGCGGTGGCGGAGCAGAGGCGCTTCTCGGTGTCCACCTTGGGC 420
QY 429 ----- 428
DB 421 AACTTGGGCGCTGGGCAAGAAGTCGCTGGAGCAGTGGGTGACCGAGGAGCGCGCTTGCCTT 480
QY 429 -----GGACGCGCCCTTTTCGCCCCAACGCTCTCTTCGACAAA 464
DB 481 TGTGCGGCTTTCGCAACACTTCGCGAGCGCCCTTTTCGCCCCAACGCTCTCTTCGACAAA 540
QY 465 GCCGTGAGCAACGTGATCGCTCCCTCACCTGCGGGCGCGCTTTCGAGTACAGACACCT 524
DB 541 GCCGTGAGCAACGTGATCGCTCCCTCACCTGCGGGCGCGCTTTCGAGTACAGACACCT 600
QY 525 CGCTTCTCAGGCTGTGGACCTAGCTCAGAGGGACTGAAGAGGAGTCTGGGCTTCTCTG 584
DB 601 CGCTTCTCAGGCTGTGGACCTAGCTCAGAGGGACTGAAGAGGAGTCTGGGCTTCTCTG 660
QY 585 CGCGAGGTGTGAATGCTGCCCCCTCTCTGTCATATCCAGCGCTGGCTGGCAAGGTC 644
DB 661 CGCGAGGTGTGAATGCTGCCCCCTCTCTGTCATATCCAGCGCTGGCTGGCAAGGTC 720
QY 645 CTACGCTTCCAAAAGGCTTTCTGACCCAGCTGGAGCTGCTAACTGACAGCAGGATG 704
DB 721 CTACGCTTCCAAAAGGCTTTCTGACCCAGCTGGAGCTGCTAACTGACAGCAGGATG 780
QY 705 ACCTGGGACCCAGCCAGCCCCCCCCGAGACTGTAAGGCGCTTCTCTGGCAGAGATGGAG 764
DB 781 ACCTGGGACCCAGCCAGCCCCCCCCGAGACTGTAAGGCGCTTCTCTGGCAGAGATGGAG 840
QY 765 AAGGCCAAGGGGAACCCCTGAGAGCAGCTTCAATGATGAGAACCTGCCCATAGTGGTGGCT 824
DB 841 AAGGCCAAGGGGAACCCCTGAGAGCAGCTTCAATGATGAGAACCTGCCCATAGTGGTGGCT 900
QY 825 GACCTGTTCTCTGCGGGAGTGTGACCACTCGACACGCTGGCTGGCGCTTCTCTGCTC 884
DB 901 GACCTGTTCTCTGCGGGAGTGTGACCACTCGACACGCTGGCTGGCGCTTCTCTGCTC 960
QY 885 ATGATCTCTACATCCGAGTGTGACGCGCGTGTCCAAAGAGAGTCTGAGCTGATAGGG 944
DB 961 ATGATCTCTACATCCGAGTGTGACGCGCGTGTCCAAAGAGAGTCTGAGCTGATAGGG 1020
QY 945 CAGGTGCGGCGACAGAGATGGGTGACAGGCTCACAATGCCCTACACCTGCCGTGATT 1004
DB 1021 CAGGTGCGGCGACAGAGATGGGTGACAGGCTCACAATGCCCTACACCTGCCGTGATT 1080
QY 1005 CATGAGTGTGACGCGCTTTTGGGACATGTCCTCCCTGGGTGTGACCCATATGACATCCCGT 1064
DB 1081 CATGAGTGTGACGCGCTTTTGGGACATGTCCTCCCTGGGTGTGACCCATATGACATCCCGT 1140
QY 1065 GACATCGAAGTACAGGCGTTCGGCATCCCTTAAGGGAACGACACTCATCACCACCTGTCA 1124
DB 1141 GACATCGAAGTACAGGCGTTCGGCATCCCTTAAGGGAACGACACTCATCACCACCTGTCA 1200
QY 1125 TCGGTGTGAGAGTATGAGCGCTCTGGGAGAACCTTTCGCTTCCACCCCGGACACTTC 1184
DB 1201 TCGGTGTGAGAGTATGAGCGCTCTGGGAGAACCTTTCGCTTCCACCCCGGACACTTC 1260
QY 1185 CTGATGCCACAGGCGCACTTTGTGAAGCGCGAGGCGCTTCTCTGCTTTTCTCAGCAGGCGC 1244
DB 1261 CTGATGCCACAGGCGCACTTTGTGAAGCGCGAGGCGCTTCTCTGCTTTCTCAGCAGGCGC 1320
QY 1245 CGTGCATGCTCGGGGAGCCCTTGGCGCGCATGAGGCTTCTCTCTTCTTCTTCTTCTTCT 1304
DB 1321 CGTGCATGCTCGGGGAGCCCTTGGCGCGCATGAGGCTTCTCTCTTCTTCTTCTTCTTCT 1380

[illegible]

FEATURES		Pfizer Products Inc. (US)	
SOURCE	Location/Qualifiers	1..1565	
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match		78.8%; Score 1211.2; DB 6; Length 1565;	
Best Local Similarity		89.5%; Pred. No. 6.8e-201;	
Matches 1403; Conservative		0; Mismatches 8; Indels 157; Gaps 3;	
QY	78	ATGGGGCTAGAACACCTGTGTCCTGCGCGTATAGTGGCCATCTTCTGCTCCTGGTG	137
DB	1	ATGGGGCTAGAACACCTGTGTCCTGCGCGTATAGTGGCCATCTTCTGCTCCTGGTG	60
QY	138	GACCTGATGACCGCGCCCAACGCTGGCTGCAAGCTACCCACAGGCCCCCTGCCACTG	197
DB	61	GACCTGATGACCGCGCCCAACGCTGGCTGCAAGCTACCCACAGGCCCCCTGCCACTG	120
QY	198	CCGGGCTGGCAACCTGCTGATGTGGACTTCCAGAACACACATACCTGCTTCGACACAG	257
DB	121	CCGGGCTGGCAACCTGCTGATGTGGACTTCCAGAACACACATACCTGCTTCGACACAG	180
QY	258	TTGGCGCGCGCTTTCGGGAGACGTTTCAGCTGCACTGGCTGAGACGCGGTGTCGTG	317
DB	181	TTGGCGCGCGCTTTCGGGAGACGTTTCAGCTGCACTGGCTGAGACGCGGTGTCGTG	240
QY	318	CTCAATGGGCTGGCGCGCTGCGGAGGCGCTGGTGACCCACCGGAGGACACCGCCGAC	377
DB	241	CTCAATGGGCTGGCGCGCTGCGGAGGCGCTGGTGACCCACCGGAGGACACCGCCGAC	300
QY	378	CGCCCGCTGTGCCATACCCAGATCTCTGGGTTTTGGGCGCGGTTCCCAA	428
DB	301	CGCCCGCTGTGCCATACCCAGATCTCTGGGTTTCGGGCGCGGTTCCCAA	360
QY	429	-----	428
DB	361	CTGGCGGCTATGGGCGCGCTGGCGGAGAGAGGCGCTTCTCCGTCCACCTTGGCG	420
QY	429	-----	428
DB	421	AAC TTGGGCTGGGCAAGATGCTGCTGGAGCATGGGTGACCGAGAGGCGCGCTTGCCTT	480
QY	429	-----GGACGCCCTTTTCGCCCCCAACGCTCTCTTGGACAAA	464
DB	481	TGTGGCGCTTTCGCCCAACCACTCCGAGCGCCCTTTCGCCCCCAACGCTCTCTTGGACAAA	540
QY	465	GCGTGAGCAACGTGATGCTCCTCTCACCTGCGGCGCGCGCTTCGAGTACGACGACCT	524
DB	541	GCGTGAGCAACGTGATGCTCCTCTCACCTGCGGCGCGCGCTTCGAGTACGACGACCT	600
QY	525	CGCTTCTCAGGCTGTGACCTAGCTCAGGAGGAGCTGAAGGAGGAGTGGGCTTTCTG	584
DB	601	CGCTTCTCAGGCTGTGACCTAGCTCAGGAGGAGCTGAAGGAGGAGTGGGCTTTCTG	660
QY	585	CGGAGGCTGTAATGCTGTCCCGCTCTCTGATATCCAGCGCTGCTGCGAGGTC	644
DB	661	CGGAGGCTGTAATGCTGTCCCGCTCTCTGATATCCAGCGCTGCTGCGAGGTC	720
QY	645	CTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGAGTGTCTAACTGAGCACAGGATG	704
DB	721	CTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGAGTGTCTAACTGAGCACAGGATG	780
QY	705	ACCTGGGACCCAGCCCGCGGAGGAGCTGAGGCGCTTCTTGGCAGAGTGGAG	764
DB	781	ACCTGGGACCCAGCCCGCGGAGGAGCTGAGGCGCTTCTTGGCAGAGTGGAG	840
QY	765	AAGGCCAAGGGAAACCTGAGAGCAGCTTCAATGATGAGAACTCGCGCATGTGGTGGCT	824
DB	841	AAGGCCAAGGGAAACCTGAGAGCAGCTTCAATGATGAGAACTCGCGCATGTGGTGGCT	900
QY	825	GACCTGTTCTCTGCGGGATGTGTGACCACTCGACCACTGCGCTGGCGCTCTCTGCTC	884

DB	901	GACCTGTTCTCTGCGGGATGTGTACCACTCGACACACCTGCGCTGGGCGCTCTCTGCTC	960
QY	885	ATGATCCTACATCCGATGTGCAGCGCGCTGTCCAAACAGAGATCGACACGCTGATAGGG	944
DB	961	ATGATCCTACATCCGATGTGCAGCGCGCTGTCCAAACAGAGATCGACACGCTGATAGGG	1020
QY	945	CAGGTGCGGCGACACAGAGATGGTGAACAGGCTCAATGCCCTVACACCACTGCGGTGATT	1004
DB	1021	CAGGTGCGGCGACACAGAGATGGTGAACAGGCTCAATGCCCTVACACCACTGCGGTGATT	1080
QY	1005	CATGAGTGCAGCGCTTTGGGGAACATGTCCTCCCTGGGTGTGACCCATATGATCCCGT	1064
DB	1081	CATGAGTGCAGCGCTTTGGGGAACATGTCCTCCCTGGGTGTGACCCATATGATCCCGT	1140
QY	1065	GACATCGAAGTACAGGGCTTCGCACTCCCTAAGGGAACGACACTCATACCAACCTGTCA	1124
DB	1141	GACATCGAAGTACAGGGCTTCGCACTCCCTAAGGGAACGACACTCATACCAACCTGTCA	1200
QY	1125	TCGGTGTGAAGGATGAGGCGCTCTGGGAGAAAGCCCTTCCGCTTCCACCCGAACTTTC	1184
DB	1201	TCGGTGTGAAGGATGAGGCGCTCTGGGAGAAAGCCCTTCCGCTTCCACCCGAACTTTC	1260
QY	1185	CTGGATGCCAGGCGCACTTTGTGAAGCGGAGGCGCTTCTGCTCTTCTCAGAGGCGCG	1244
DB	1261	CTGGATGCCAGGCGCACTTTGTGAAGCGGAGGCGCTTCTGCTCTTCTCAGAGGCGCG	1320
QY	1245	CSTGATGCTCGGGGAGCGCTGCGCGCATGAGAGCTTCTTCTTCTTCTTCTTCTTCTTCT	1304
DB	1321	CSTGATGCTCGGGGAGCGCTGCGCGCATGAGAGCTTCTTCTTCTTCTTCTTCTTCTTCT	1380
QY	1305	CTGAGCACTTTCAGCTTCTCGGTGCGCCACTGAGAGCGCGCGCGCGCGCGCGCGCGCG	1364
DB	1381	CTGAGCACTTTCAGCTTCTCGGTGCGCCACTGAGAGCGCGCGCGCGCGCGCGCGCGCG	1440
QY	1365	TTTGTCTTCTGAGTGCAGCCATCCCTCTATGAG-CTTTGTGCTGTGCGCGCGCGCGCGCG	1423
DB	1441	TTTGTCTTCTGAGTGCAGCCATCCCTCTATGAGTATTTGTGCTGTGCGCGCGCGCGCGCG	1500
QY	1424	GSTACTAGTCCCGAGCGCTGCTCCCTAGCAGAGGCTCTAATGTACAAATAAGCAATGTG	1483
DB	1501	GSTACTAGTCCCGAGCGCTGCTCCCTAGCAGAGGCTCTAATGTACAAATAAGCAATGTG	1557
QY	1484	GTAGTTCC 1491	
DB	1558	GTAGTTCC 1565	
RESULT 9			
AR084374			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
/organism="unknown"			
/mol_type="unassigned DNA"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1392; Conservative			
QY	78	ATGGGGCTAGAACACCTGTGTCCTGCGCGTATAGTGGCCATCTTCTGCTCCTGGTG	137

|||||
1 ATGGGGCTAGAAACACTGTGGTCCCTCGCGGTGATAGTGGCCATCTTCTGCTCTCTGTG 60
QY 138 GACCTGATGACACCGGGGCAACGCTGGGTGACACCTACTACACAGGCCCTCTGCACTG 197
Db 61 GACCTGATGACACCGGGGCAACGCTGGGTGACACCTACTACACAGGCCCTCTGCACTG 120
QY 198 CCCGGCTGGGCAACCTGCTGATCTGGACTTCCAGAAACACACCATATCTGTTCCAGCAG 257
Db 121 CCCGGCTGGGCAACCTGCTGATCTGGACTTCCAGAAACACACCATATCTGTTCCAGCAG 180
QY 258 TTGCGGCGCGCTTCCGGGACGCTTTACGCTCTGACGCTGGCTGACGCGCGGTGCTGCTG 317
Db 181 TTGCGGCGCGCTTCCGGGACGCTTTACGCTCTGACGCTGGCTGACGCGCGGTGCTGCTG 240
QY 318 CTCAATGGGCTGGCGGCGCTGGCGAGGGCTGTGTGACCCAGCGGAGGACACCGCCGAC 377
Db 241 CTCAATGGGCTGGCGGCGCTGGCGAGGCGCTGTGTGACCCAGCGGAGGACACCGCCGAC 300
QY 378 CGCCGCGCTGTGCCATCACCCAGATCTCGGTTTTTGGGCGCGCTTCCCAA----- 428
Db 301 CGCCGCGCTGTGCCATCACCCAGATCTCGGTTTTTGGGCGCGCTTCCCAAAGGGGTGTTCC 360
QY 429 ----- 428
Db 361 CTGGCGCGCTATGGGCGCGGTGGCGGAGAGAGCGCTTCTCGTGTCCACCTTTGGCG 420
QY 429 ----- 428
Db 421 AACTTTGGGCTGGGCAAGAACTGCTGGAGCAGTGGGTGACCGAGGAGCGCGCTTGCCTT 480
QY 429 -----GGAGCGCCCTTTTCGCCCCCAAGCTCTCTTGACAAA 464
Db 481 TGTGGCGCTTCCGCAACCACTCCGAGCGCCCTTTTCGCCCCCAAGCTCTCTTGACAAA 540
QY 465 GCGGTGAGCAACGCTGATCCCTTCCCTCACCTGCGGGCGCGCTTCCGAGTACGACGACCT 524
Db 541 GCGGTGAGCAACGCTGATCCCTTCCCTCACCTGCGGGCGCGCTTCCGAGTACGACGACCT 600
QY 525 CGCTTCTCAGGCTCTGACCTAGCTCAGAGGAGCTGAAGAGGAGTCTCGGCTTTCTG 584
Db 601 CGCTTCTCAGGCTCTGACCTAGCTCAGAGGAGCTGAAGAGGAGTCTCGGCTTTCTG 660
QY 585 CGCGAGGTGCTGAATGCTTCCCGTCTCCCTGCTGATATCCAGCGCTGGCTGGCAAGTGC 644
Db 661 CGCGAGGTGCTGAATGCTTCCCGTCTCCCTGCTGATATCCAGCGCTGGCTGGCAAGTGC 720
QY 645 CTACGCTTCCAAAAGGCTTTCTGACCCAGCTGGATGAGCTGAACCTGAGCACAGGATG 704
Db 721 CTACGCTTCCAAAAGGCTTTCTGACCCAGCTGGATGAGCTGAACCTGAGCACAGGATG 780
QY 705 ACCTGGGACCCAGCCAGCCCGCCCGAGACCTGACTGAGGCTTCTCTGGCAGAGATGAG 764
Db 781 ACCTGGGACCCAGCCAGCCCGCCCGAGACCTGACTGAGGCTTCTCTGGCAGAGATGAG 840
QY 765 AAGGCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGGTGGCT 824
Db 841 AAGGCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGGTGGCT 900
QY 825 GACCTGTTCTCTGCGGGATGGTGACCACTCGACACGCTGGCTGGGCGCTCTCTGCTC 884
Db 901 GACCTGTTCTCTGCGGGATGGTGACCACTCGACACGCTGGCTGGGCGCTCTCTGCTC 960
QY 885 ATGATCTTACATCCGGATGTGACGCGCGCTGTGCAACAGGAGATCGACGAGCTGATAGGG 944
Db 961 ATGATCTTACATCCGGATGTGCA---TCGCTGTGCAACAGGAGATCGACGAGCTGATAGGG 1018
QY 945 CAGGTGCGGCGACAGAGATGGGTGACACAGGCTCAATGCGCTACACATGCGCGTGATT 1004
Db 1019 CAGGTGCGGCGACAGAGATGGGTGACACAGGCTCAATGCGCTACACATGCGCGTGATT 1078
QY 1005 CATGAGGTGACGCGCTTTTGGGAGCATCTGCCCTGGGTGTGACCCATATGACATCCCGT 1064
|||||

Db 1079 CACGAGTGCAGCGCTTTTGGGACATCATCCCTCGAGTGTGACCCATATGACATCCCGT 1138
QY 1065 GACATCAAGTACAGGGCTTCCGATCCCTAAGGAAGACACTCATCACCAACTGTCA 1124
Db 1139 GACATCAAGTACAGGGCTTCCGATCCCTAAGGAAGACACTCATCACCAACTGTCA 1198
QY 1125 TCGGTGCTGAAGGATGAGGCGCTCTGGAGAAAGCCCTTTCGGCTTCCACCCCGAAACACTTC 1184
Db 1199 TCGGTGCTGAAGGATGAGGCGCTCTGGAGAAAGCCCTTTCGGCTTCCACCCCGAAACACTTC 1258
QY 1185 CTGATGCCAGGCGCACTTTGTGAAGCGAGAGCCCTTCTGCGCTTTCTCAGCAGGCGC 1244
Db 1259 CTGATGCCAGGCGCACTTTGTGAAGCGAGAGCCCTTCTGCGCTTTCTCAGCAGGCGC 1318
QY 1245 CGTCATGCTCGGGGAGCGCTTGGCCCGCATGAGCTCTTCTCTTTCACCTCCCTG 1304
Db 1319 CGTCATGCTCGGGGAGCGCTTGGCCCGCATGAGCTCTTCTCTTTCACCTCCCTG 1378
QY 1305 CTGCAGCACTTTCAGCTTCTCGGTGCCACTGGAGAGCGCCCGGCCAGCACCATGGTGTG 1364
Db 1379 CTGCAGCACTTTCAGCTTCTCGGTGCCCGCGAGAGCGCCCGGCCAGCACCTCTCGTGTG 1438
QY 1365 TTTGCTTTCTCTGCTGACCCCATCCCTATGAGCTTTTGTGTGTGCGCCCGCTAGATGGG 1424
Db 1439 GTCAAGCTTCTGCTGTGACCCCATCCCTACGAGCTTTGTGTGTGCGCCCGCTAGATGGG 1498
QY 1425 GTACTAGTCCCGAGCTGTCTCCCTAGCAGAGCTCTAATGTACATAAGCAATGTGG 1484
Db 1499 GTACTAGTCCCGAGCTGTCTCCCTAGCAGAGCTCTAATGTACATAAGCAATGTGG 1558
QY 1485 TAGTTCCA 1492
Db 1559 TAGTTCCA 1566

RESULT 10

E15820
LOCUS E15820 1545 bp DNA linear PAT 28-JUL-1999
DEFINITION cDNA encoding human cytochrome P450-2d6.
ACCESSION E15820
VERSION E15820.1 GI:5710503
KEYWORDS JP 1998101698-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1545)
AUTHORS Narita,Y., Sawaguchi,S., Lari,W.H., Hisakawa,Y., Soma,G., Kako,M.
and Miyagawa,H.
LIVER/KIDNEY MICRO SOME OCCURRING IN AUTOIMMUNOHEPATITIS TYPE II
(AII TYPE II)
JOURNAL Patent: JP 1998101698-A 1 21-APR-1998;
HOKEN KAGAKU KENKYUSHO:KK
OS Homo sapiens (human)
COMMENT PN JP 1998101698-A/1
PD 21-APR-1998
PF 24-SEP-1996 JP 1996273055
PI NARITA YOICHI, SAWAGUCHI SHICHIRO, LARI W HANKINS, PI
HISAKAWA YOSHIZO,
PI SOMA GENICHIRO, KAKO MAKOTO, MIYAGAWA HIROSHI PC
C07K14/47, C07H21/04, C07K16/18, C12N15/09, C12Q1/68, G01N33/53, PC
G01N33/564, C12P21/02, (C12P21/02, C12R1:19);
PC G01N33/576//C12P21/02, (C12P21/02, C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
Key
FH
FT 1..1545
FT /organism='Homo sapiens',
FT /tissue type='liver',
FT 1..1494
FT CDS


```
FT misc_RNA 847. /product='human cytochrome P450-2d6', FT
FT 1431 /product='antigen for AIH Type2a LKM1 antibody
FT and for AIH Type2b LKM1 antibody'
FT 349..825 /product='antigen for AIH Type2a LKM1 FT
FT antibody'
FEATURES
source Location/Qualifiers
1..1545
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 78.1%; Score 1201; DB 6; Length 1545;
Best Local Similarity 88.9%; Pred. No. 4.1e-199;
Matches 1373; Conservative 0; Mismatches 19; Indels 153; Gaps 1;
QY 78 ATGGGGCTAGAGCACTGGTGGCCCTGGCGGTGATGATGGGCATCTTCTGCTCTGGTG 137
Db 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCGCATGATGATGGGCATCTTCTGCTCTGGTG 60
QY 138 GACCTGATGACACGGCGCAACGCTGGCTGCACTACTACCCAGGCCCTGCCACTG 197
Db 61 GACCTGATGACACGGCGCAACGCTGGCTGCACTACTACCCAGGCCCTGCCACTG 120
QY 198 CCGGGCTGGCAACCTGCTGATGATGGACTTCCAGAACACACCACTACTGCTTCGACCA 257
Db 121 CCGGGCTGGCAACCTGCTGATGATGGACTTCCAGAACACACCACTACTGCTTCGACCA 180
QY 258 TTGCGGCGCGCTTGGGGAACGTTTACGCTGCACTGGCTGGCGCTGGAGCGCGTGGTG 317
Db 181 TTGCGGCGCGCTTGGGGAACGTTTACGCTGCACTGGCTGGCGCTGGAGCGCGTGGTG 240
QY 318 CTCATGGCTGGCGCGCTGGCGAGCGCTGGTGAACCCAGCGGAGGACACCGCGAC 377
Db 241 CTCATGGCTGGCGCGCTGGCGAGCGCTGGTGAACCCAGCGGAGGACACCGCGAC 300
QY 378 CGCGCGCTGTGCCATACCCAGATCCCTGGGTTTGGGCGCGCTTCCCAA----- 428
Db 301 CGCGCGCTGTGCCATACCCAGATCCCTGGGTTTGGGCGCGCTTCCCAAAGGGTGTTC 360
QY 429 ----- 428
Db 361 CTGGCGCGCTATGGGCGCGCTGGCGAGCAGAGCGGCTTCTCGCTCTCCACTTGGCG 420
QY 429 ----- 428
Db 421 AACTTGGGCTGGGCAAGTCTGCTGGAGCAGTGGTGACCGGAGGCGCGCTTGCCTT 480
QY 429 -----GGACCGCCCTTTGGCCCAACCGTCTCTTGGACAAA 464
Db 481 TGTGGCGCTTGGCCAAACACTCCGGAGCGCCCTTTGGCCCAACCGTCTTGGACAAA 540
QY 465 GCGGTGAGCAACGTGATGCGCTTCTTCACTCGCGGCGCGCTTCGAGTACGACACCGCT 524
Db 541 GCGGTGAGCAACGTGATGCGCTTCTTCACTCGCGGCGCGCTTCGAGTACGACACCGCT 600
QY 525 CGCTTCTCAGGCTGTGACCTAGCTCAGGAGGACTGAAGGAGAGTCTGGGCTTCTTG 584
Db 601 CGCTTCTCAGGCTGTGACCTAGCTCAGGAGGACTGAAGGAGAGTCTGGGCTTCTTG 660
QY 585 CGCGAGGTGCTGAAGTCTGCTCCCTCTCTGATATCCAGCGCTGGCTGGCAGGCTC 644
Db 661 CGCGAGGTGCTGAAGTCTGCTCCCTCTCTGATATCCAGCGCTGGCTGGCAGGCTC 720
QY 645 CTACGCTTCCAAAGGCTTCTTCAACCGAGCTGGAGTCTTCACTGAGCAGGATG 704
Db 721 CTACGCTTCCAAAGGCTTCTTCAACCGAGCTGGAGTCTTCACTGAGCAGGATG 780
QY 705 ACCTGGGACCGACCGCGCGCGCGAGACTGAGCGCTTCTTGGCAGAGATGAG 764
|||||
```

```
Db 781 ACCTGGGACCCAGCCAGCCGAGCCCCCGAGACCTGACTGAGGCTTCTTGGCAGAGATGGAG 840
QY 765 AAGGCCAAGGGGAAACCTCGAGAGCAGCTTCAATGATGAGAACTTGGCGCATAGTGGTGGCT 824
|||||
Db 841 AAGGCCAAGGGGAAACCTCGAGAGCAGCTTCAATGATGAGAACTTGGCGCATAGTGGTGGCT 900
|||||
QY 825 GACCTGTTCTCTGCGGGGATGATGACCACTCGACCACTCGACCACTCGACCACTCGACCACT 884
|||||
Db 901 GACCTGTTCTCTGCGGGATGATGATGACCACTCGACCACTCGACCACTCGACCACTCGACCACT 960
|||||
QY 885 ATGATCTTACATCCCGATGTCAGCGCGCTGTCAGAGGAGATGAGAGGATGAGAGGATGAGAGG 944
|||||
Db 961 ATGATCTTACATCCCGATGTCAGCGCGCTGTCAGAGGAGATGAGAGGATGAGAGGATGAGAGG 1020
|||||
QY 945 CAGGTGCGCGGACACAGAGATGGGTGACAGGCTCAGATGCGCTTACACCACTGCGCGTGAAT 1004
|||||
Db 1021 CAGGTGCGCGGACACAGAGATGGGTGACAGGCTCAGATGCGCTTACACCACTGCGCGTGAAT 1080
|||||
QY 1005 CATGAGGTGACAGCGCTTTGGGGAACATCGTTCCTTGGGTGTCACCATATGACATCCCGT 1064
|||||
Db 1081 CATGAGGTGACAGCGCTTTGGGGAACATCGTTCCTTGGGTGTCACCATATGACATCCCGT 1140
|||||
QY 1065 GACATCGAATACAGGGCTTCGGCATCCCTAAGGGAACGACACTATACCAACCTGTCA 1124
|||||
Db 1141 GACATCGAATACAGGGCTTCGGCATCCCTAAGGGAACGACACTATACCAACCTGTCA 1200
|||||
QY 1125 TCGGTGCTGAAGGATGAGCGCTTGGGAGAGCCCTTCCGCTTCCACCCGAACTTTC 1184
|||||
Db 1201 TCGGTGCTGAAGGATGAGCGCTTGGGAGAGCCCTTCCGCTTCCACCCGAACTTTC 1260
|||||
QY 1185 TTGGATGCCAGGGCCACTTTTGAAGCGGAGGCGCTTCTGCTCTTCTCAGCAGGCGCG 1244
|||||
Db 1261 TTGGATGCCAGGGCCACTTTTGAAGCGGAGGCGCTTCTGCTCTTCTCAGCAGGCGCG 1320
|||||
QY 1245 CGTCATGCTTGGGAGCGCGCTTGGCGCGCATGAGAGCTTCTTCTTCTTCACTCCCTG 1304
|||||
Db 1321 CGTCATGCTTGGGAGCGCGCTTGGCGCGCATGAGAGCTTCTTCTTCTTCACTCCCTG 1380
|||||
QY 1305 CTGACGACTTTCAGCTTCTCGGTGCGCCACTGACAGCGCGCGCGCGCGCGCGCGCGCG 1364
|||||
Db 1381 CTGACGACTTTCAGCTTCTCGGTGCGCCACTGACAGCGCGCGCGCGCGCGCGCGCGCG 1440
|||||
QY 1365 TTGCTTCTTCTGAGACCCCATCCCTTATGAGCTTGTGCTGTCGCCCGCTTAGAATGGG 1424
|||||
Db 1441 TTGCTTCTTCTGAGACCCCATCCCTTATGAGCTTGTGCTGTCGCCCGCTTAGAATGGG 1500
|||||
QY 1425 GTACCTAGTCCCGAGCGCTGCTCCCTAGCCAGAGGCTCTAATGTAC 1469
|||||
Db 1501 GTACCTAGTCCCGAGCGCTGCTCCCTAGCCAGAGGCTCTAATGTAC 1545
|||||
RESULT 11
AF301911
LOCUS AF301911 1585 bp mRNA linear PRI 21-JAN-2003
DEFINITION Macaca fuscata cytochrome P450 2D mRNA, complete cds.
ACCESSION AF301911
VERSION AF301911.1 GI:18481429
KEYWORDS
SOURCE Macaca fuscata (Japanese macaque)
ORGANISM Macaca fuscata
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Macaca.
REFERENCE 1 (bases 1 to 1585)
AUTHORS Hichiya,H., Takemi,C., Tsuzuki,D., Yamamoto,S., Asaoka,K.,
Suzuki,S., Satoh,T., Shinoda,S., Kataoka,H. and Narimatsu,S.
TITLE Complementary DNA cloning and characterization of cytochrome P450
JOURNAL 2D9 from Japanese monkey liver
MEDLINE Biochem. Pharmacol. 64 (7), 1101-1110 (2002)
PUBMED 22219831
PUBMED 12234613
REFERENCE 2 (bases 1 to 1585)
AUTHORS Hichiya,H., Yamamoto,S., Asaoka,K. and Narimatsu,S.
```

TITLE Direct Submission
JOURNAL Submitted (30-AUG-2000) Laboratory of Health Chemistry, Faculty of
Pharmaceutical Sciences, Okayama University, 1-1-1 Tsushima-naka,
Okayama 700-8530, Japan
FEATURES Location/Qualifiers
source 1..1585
/organism="Macaca fuscata"
/mol_type="mRNA"
/db_xref="taxon:9542"
/tissue_type="liver"
41..1534
/note="CYP2D"
/codon_start=1
/product="cytochrome P450 2D"
/protein_id="AA173443.1"
/db_xref="GI:18481430"
/translation="MQLDALVPLAVTVAIILLVLMHRRQRAARYPPGLPLPFLFGL
KLLHVDFOFTYCFQELRRFQDFSLQAWTVVLLNGLAAVAEALVTHGEDTADRP
VPVITQIILGFGPRSGVFLARYGASWRORRFSVTLNGLGKSGTOWTGTGAACL
CAAPANHSRPFPRNGLLDKAVSNVIALITCGRRPEYDDPRFLRLDLAORGLKEEPG
FLREVNAPVLLHLPALAGKVRQKAPLQDLDELTEHRMTWDPAPRDLTRVQL
AMEKAKGNPSSFNENLRIIVADLFSAGVMTSTLLAWGLLMLILPDDVORRVQOE
IDDVIGVYRPALEGAHMPYTPAVKEVQRFQDFVLDQGHFVDEALPFSAGRACLGELPLAR
TTLITNLSSVLKDEAVWEKPFHPEHFLDQGHFVDEALPFSAGRACLGELPLAR
MELFLFTSLQLHRSFVPTGQPRSHHGVAFLVSPSPYELCAVPR"

CDS

ORIGIN

Query Match 76.9%; Score 1181.4; DB 9; Length 1585;
Best Local Similarity 87.6%; Pred. No. 1.1e-195;
Matches 1391; Conservative 0; Mismatches 41; Indels 155; Gaps 3;
QY 38 CAGGTGTCTCCAGAGAGCCCATTTGGTAGTCAGGAGTATGGGGCTAGAGCACTGGT 97
DB 1 CGGGGGGTGTCCAGAGAGCCCGAGGGGGAGTAAGGACAGCAATGGAGCTAGATGCACCTGGT 60
QY 98 GCCCTGGCCGTGTAGTGGCCATCTTCTGCTCTCTGGTGGACCTGTGATGCACCGCGCCA 157
DB 61 GCCCTGGCTGTGACAGTGGCCATCATCTGCTCTCTGGTGGACCTGTGATGCACCGCGCCA 120
QY 158 AGCTGGGTGTGACGCTACTACAGAGCCCTCTGCACTGCGCGGGCTGGGCAACTGCT 217
DB 121 AGCTGGGGCGGACGCTACCCAGGCGCCCTGCACTGCGCCCTTCTGGGCAAACTGCT 180
QY 218 GCATGTGACTTCCAGACACACCATCTGTTCCAGCAGTTCGCGCGCGCTTCGGGGA 277
DB 181 GCATGTGACTTCCAGACACACCATCTGTTCCAGCAGTTCGCGCGCGCTTCGGGGA 240
QY 278 CGTGTTCAGCTGCGAGTGGCTGACGCGCGGTGCTGCTCAATGGGCTGGCGCGCT 337
DB 241 CGTGTTCAGCTGCGAGTGGCTGACGCGCGGTGCTGCTCAATGGGCTGGCGCGCT 300
QY 338 GCGGAGGCGGTGTGACCCAGCGAGAGACACCGCGAGCCCGCTGCTGCGCCATCAC 397
DB 301 GCGGAGGCGGTGTGACCCAGCGAGAGACACCGCGAGCCCGCTGCTGCGCCATCAC 360
QY 398 CCAGATCTGGGTTTGGGCGCGCTTCCAA----- 428
DB 361 CCAGATCTGGGTTTGGGCGCGCTTCCAAAGGGGTGTTCTGGCGCGCTATGGCGCTC 420
QY 429 ----- 428
DB 421 GTGGCGGAGCAGAGCGGCTTCTCGCTCTCCACTTGGCAACTTGGGCGCTGGGCAAGAA 480
QY 429 ----- 428
DB 481 GTGGCGGAGCAGTGGGTGACCGGGAGGCGCGCTTGTGCGCGCTTCGCGCAACCA 540
QY 429 -----GGAGCGCCCTTTCGCCCAACCGTCTCTTGGACAAACCGGTGACCAAGTATCGC 484
DB 541 CTCGCGAGCGCCCTTTCGCCCAACCGTCTCTTGGACAAACCGGTGACCAAGTATCGC 600
QY 485 CTCCTCTACCTGCGGCGCGCTTCGAGTACGACGACCTCGCTTCTCCTCAGCTCTCGGA 544

DB 601 CTCCTCACCTGCGGGCGCGCTTTCGAGTACGACGACCCCTCGCTTCTCTCAGGCTGCTGGA 660
QY 545 CCTAGCTCAGAGAGGACTGAAGAGGAGTCCGGCTTTCGCGGAGGTGCTGAATGCTGT 604
DB 661 CCTAGCTCAGAGAGGACTGAAGAGGAGGCGGGCTTTCGCGGAGGTGCTGAATGCTGT 720
QY 605 CCGCTCTCTCTCATATCCAGCGCTGGCTGGCAAGGTCTCTAGCTTCCAAAAGGCTTT 664
DB 721 CCGCTCTCTCTCATATCCAGCGCTGGCTGGCAAGGTCTCTAGCTTCCAAAAGGCTTT 780
QY 665 CCTCACCGAGCTGAGTGTCTAACTGAGCAGAGTACCTGGGAGCCAGGCCAGCC 724
DB 781 CCGTACCGAGCTGAGTGTCTAACTGAGCAGAGTACCTGGGAGCCAGGCCAGCC 840
QY 725 CCGCGGAGACCTGACTGAGGCTTCTGGGAGAGATGGAGAGGCCAAGGGGAACCTGGA 784
DB 841 CCGCGGAGACCTGACTGAGGCTTCTGGGAGAGATGGAGAGGCCAAGGGGAACCTGGA 900
QY 785 GAGCAGCTTCAATGATGAGAACCTCGCATAGTGGTGGCTGACCTGTTCTCTGCGGGAT 844
DB 901 GAGCAGCTTCAATGATGAGAACCTCGCATAGTGGTGGCTGACCTGTTCTCTGCGGGAT 960
QY 845 GGTGACCACTCGACACCGCTGGGCTCTGCTGCTCATATGATCTTACATCCGAGTGT 904
DB 961 GGTGACCACTCGACACCGCTGGGCTCTGCTGCTCATATGATCTTACATCCGAGTGT 1020
QY 905 GCAGGCGGTGTCCAAAGAGGATCGAGAGCTGATAGGCGAGGTGGCGGACACAGAGAT 964
DB 1021 GCAGGCGGTGTCCAAAGAGGATCGAGAGCTGATAGGCGGTGTGGCGGTCTCTCGGCT 1080
QY 965 GGGTGACCGAGCTCACATGCCCTACACCTGCGGTGATTTCATGAGGTGACGCGCTTGG 1024
DB 1081 GGGTGAAAGAGCTCACATGCCCTACACCTGCGGTGATTTCATGAGGTGACGCGCTTGG 1140
QY 1025 GGACATCGTCCCTGGGTGTGACCCATATGACATCCGCTGACATCGAAGTACAGGGCTT 1084
DB 1141 GGACATCGTCCCTGGGTGTGACCCATATGACATCCGCTGACATCGAAGTACAGGGCTT 1200
QY 1085 CCGCATCCCTTAAGGAGAGGACACTCATCAACCTGTCATCGGTGCTGAAGGATGAGGC 1144
DB 1201 CCGCATCCCTTAAGGAGAGGACACTCATCAACCTGTCATCGGTGCTGAAGGATGAGGC 1260
QY 1145 CGTCTGGAGAGAGCCCTTCCGCTTCCACCCCGAACAACCTTCTGGATGCCAGGCGCACTT 1204
DB 1261 CGTCTGGAGAGAGCCCTTCCGCTTCCACCCCGAACAACCTTCTGGATGCCAGGCGCACTT 1320
QY 1205 TGTGAAGCGGAGGCTTCTGCTGCTTCTCAGCAGGCGCGCTGATGCTCTCGGGAGCC 1264
DB 1321 TGTGAAGCGGAGGCTTCTGCTGCTTCTCAGCAGGCGCGCTGATGCTCTCGGGAGCC 1380
QY 1265 CCGCGCGGATGAGGCTTCTCTTCTTCACTTCTGCTGCTGAGCACTTTCAGCTTCTC 1324
DB 1381 CCGCGCGGATGAGGCTTCTCTTCTTCACTTCTGCTGCTGAGCACTTTCAGCTTCTC 1440
QY 1325 GGTCCCACTGGACAGCCCGCGCGCAGCACCATGTGTCTTGTCTTCTTCTTCTGAGCC 1384
DB 1441 GGTCCCACTGGACAGCCCGCGCGCAGCACCATGTGTCTTGTCTTCTTCTTCTGAGCC 1500
QY 1385 ATCCCGCTATGAGCTTGTGCTGCTGCGCGCTAGAGTGGGCTACCTAGTCCCGGCTG 1444
DB 1501 ATCCCGCTATGAGCTTGTGCTGCTGCGCGCTAGAGTGGGCTAGAGTGGGCTAGAGT 1558
QY 1445 TCCCTAGCAGAGGCTCTAATGTACAA 1471
DB 1559 TCCCTAGCAGAGGCGCTGATGTACAA 1585

RESULT 12
E10868
LOCUS E10868 1494 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding human cytochrome P450.
ACCESSION E10868
VERSION E10868.1 GI:22027962

DEFINITION Sequence 17 from patent US 6620593.
ACCESSION AR399336
VERSION AR399336.1 GI:40141144
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1494)
HAYASHI, K.; SAKAKI, T.; YABUSAKI, Y.; KOMAI, K.; KANEKO, H. and
NAKATSUKA, I.
TITLE Method for safety evaluation of chemical compound using recombinant
yeast expressing human cytochrome P450
JOURNAL Patent: US 6620593-A 17 16-SEP-2003;
FEATURES Location/Qualifiers
source 1..1494
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 76.3%; Score 1173.2; DB 6; Length 1494;
Best Local Similarity 89.6%; Pred. No. 2.8e-194;
Matches 1338; Conservative 0; Mismatches 3; Indels 153; Gaps 1;
QY 78 ATGGGGCTAGAGCACTGGTGGCCCTGGCGCTGATGAGTGGCCATCTTCTGCTCTCTGGTG 137
DB 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCGCTGATGAGTGGCCATCTTCTGCTCTCTGGTG 60
QY 138 GACCTGATGACACCGCGGCAACGCTGGGCTGACGCTACTACACAGGCCCCCTGCACTG 197
DB 61 GACCTGATGACACCGCGGCAACGCTGGGCTGACGCTACTACACAGGCCCCCTGCACTG 120
QY 198 CCGGGCTGGGCAACCTGCTGATGCTGACCTTCCAGACACACCATGCTTCGACACG 257
DB 121 CCGGGCTGGGCAACCTGCTGATGCTGACCTTCCAGACACACCATGCTTCGACACG 180
QY 258 TTGCGGCGCGCTTCCGGGACGCTGCTGACCTGACGCTGGCTGACGCTGGCTGCTGCTG 317
DB 181 TTGCGGCGCGCTTCCGGGACGCTGCTGACCTGACGCTGGCTGACGCTGGCTGCTGCTG 240
QY 318 CTCATGGGCTGGCGGCGCTGGCGAGGCTGGTGGACGACGCGGAGGACACCGCGAC 377
DB 241 CTCATGGGCTGGCGGCGCTGGCGAGGCTGGTGGACGACGCGGAGGACACCGCGAC 300
QY 378 CGCCGCGCTGGCCATCACCAGATCTGGTGGTGGTGGGCGCGCTCCCAA----- 428
DB 301 CGCCGCGCTGGCCATCACCAGATCTGGTGGTGGTGGGCGCGCTCCCAAAGGGGTGTTTC 360
QY 429 ----- 428
DB 361 CTGGCGCGCTATGGCGCGCGTGGCGCGAGGAGGCGCTTCTCGTCTCCACCTTGGCG 420
QY 429 ----- 428
DB 421 AACTTGGGCTGGGCAAGAAGTCGCTGGGAGAGTGGGTGACCGAGGCGCGCTTGCCTT 480
QY 429 -----CGAGCGCCCTTTTCGCCCAACGCTCTCTTGACAAA 464
DB 481 TGTGGCGCTTGGCCAAACCACTCCGGAGCGCCCTTTGCCCAACGCTCTCTTGACAAA 540
QY 465 GCGGTGAGCAACGCTGATCGCTTCCCTCACCTGCGGCGCGCTTCCAGTAGACACACCT 524
DB 541 GCGGTGAGCAACGCTGATCGCTTCCCTCACCTGCGGCGCGCTTCCAGTAGACACACCT 600
QY 525 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGACTGAAGAGGAGTGGGCTTTCTG 584
DB 601 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGACTGAAGAGGAGTGGGCTTTCTG 660
QY 585 CGCGAGGTGCTGAATGCTGCTCCCGCTCTCTGCAATATCCAGCGCTGGTGGCAAGTTC 644
DB 661 CGCGAGGTGCTGAATGCTGCTCCCGCTCTCTGCAATATCCAGCGCTGGTGGCAAGTTC 720
QY 645 CTACCGCTTCCAAAAGGCTTCTCTGACCCAGCTGGATGAGTCTAACTGAGCACAGGATG 704

DB 721 CTACGCTTCCAAAAGGCTTCTCTGACCCAGCTGGATGAGTGTCTAACTGACGAGGATG 780
QY 705 ACCTGGGACCCAGCCAGCCCGGAGACCTGACTGAGGCTTCTTGGCAGAGATGAG 764
DB 781 ACCTGGGACCCAGCCAGCCCGGAGACCTGACTGAGGCTTCTTGGCAGAGATGAG 840
QY 765 AAGGCCAAGGGGAACCTCTGAGAGCAGCTTCAATGATGAGAACCTCGCATAGTGGTGGCT 824
DB 841 AAGGCCAAGGGGAACCTCTGAGAGCAGCTTCAATGATGAGAACCTCGCATAGTGGTGGCT 900
QY 825 GACCTGTTCTCTGCGGAGTGGTACCACTCGACGCTGAGGCTGGGGCTCTCTGCTC 884
DB 901 GACCTGTTCTCTGCGGAGTGGTACCACTCGACGCTGGGCTGGGGCTCTCTGCTC 960
QY 885 ATGATCTACATCCGATGTCAGGCGCTGTCACACGAGGATCGAGCGTAGAGG 944
DB 961 ATGATCTACATCCGATGTCAGGCGCTGTCACACGAGGATCGAGCGTAGAGG 1020
QY 945 CAGGTGGCGGACCCAGAGATGGGTGACGAGCTCACATGCCCTACCACTGCCGTGATT 1004
DB 1021 CAGGTGGCGGACCCAGAGATGGGTGACGAGCTCACATGCCCTACCACTGCCGTGATT 1080
QY 1005 CATGAGGTGACGCGCTTTTGGGACATCGTCCCTGGGTGTGACCCATATGACATCCCGT 1064
DB 1081 CATGAGGTGACGCGCTTTTGGGACATCGTCCCTGGGTGTGACCCATATGACATCCCGT 1140
QY 1065 GACATCGAGTACAGGCTTCCGATCCCTAAGGGAAGCAGCACTCATCAACCACTGTCA 1124
DB 1141 GACATCGAGTACAGGCTTCCGATCCCTAAGGGAAGCAGCACTCATCAACCACTGTCA 1200
QY 1125 TCGGTGCTGAAGGATGAGGCGCTCTGGGAGAACGCTTCCGCTTCCACCCCGAAGCACTTC 1184
DB 1201 TCGGTGCTGAAGGATGAGGCGCTCTGGGAGAACGCTTCCGCTTCCACCCCGAAGCACTTC 1260
QY 1185 CTGATGCCAGGCGCACTTTGTGAAGCGGAGGCGCTTCTGCTTTCTAGAGGCGCGC 1244
DB 1261 CTGATGCCAGGCGCACTTTGTGAAGCGGAGGCGCTTCTGCTTTCTAGAGGCGCGC 1320
QY 1245 CGTCATGCTTCGGGAGCGCTGCGCGGATGAGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCT 1304
DB 1321 CGTCATGCTTCGGGAGCGCTTGGCGGATGAGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCT 1380
QY 1305 CTGAGCACTTTCAGTCTTCTGCTGCCACTGACAGCGCCCGGCGGAGCAGCAGCAGTGGTGT 1364
DB 1381 CTGAGCACTTTCAGTCTTCTGCTGCCACTGACAGCGCCCGGCGGAGCAGCAGCAGTGGTGT 1440
QY 1365 TTTGCTTCTTCTGTTGACCCCATCCCTATGAGCTTTGTGTGTGCTGCTGCTGCTGCTGCTG 1418
DB 1441 TTTGCTTCTTCTGTTGACCCCATCCCTATGAGCTTTGTGTGTGCTGCTGCTGCTGCTGCTG 1494
RESULT 14
LOCUS E10647 1494 bp RNA linear PAT 29-SEP-1997
DEFINITION Human cDNA encoding cytochrome P4502D6.
ACCESSION E10647
VERSION E10647.1 GI:22027703
KEYWORDS JP 1996027199-A/8.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1494)
AUTHORS Funae, Y., Imaka, S., Matsuki, Y., Hayashi, K. and Yabusaki, Y.
TITLE ANTIBODY RECOGNIZING CYTOCHROME P4502D6 ORIGINATED FROM MAN
JOURNAL Patent: JP 1996027199-A 8 30-JAN-1996;
SUMITOMO CHEM CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1996027199-A/8
PD 30-JAN-1996
PI 15-JUL-1994 JP 1994164186
PI FUNAE YOSHIIKO, IMAOKA SUSUMU, MATSUKI YASUSHI, HAYASHI KOJI,
PI YABUSAKI YOSHIYASU

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 01:39:30 ; Search time 354.102 Seconds:
(without alignments)
18439.552 Million cell updates/sec

Title: US-09-820-788A-1
Perfect score: 1537
Sequence: 1 cctgctgtctctgtgcc.....aaaaaaaaaaaaaaaaaaaaa 1537

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1537	100.0	1537	7 AAL53565	Aal53565 cDNA enco
2	1241.4	80.8	1567	7 ADB25779	Adb25779 Human CYP
3	1241.4	80.8	1567	7 ADB25777	Adb25777 Human CYP
4	1241.4	80.8	1567	7 ADB25776	Adb25776 Human CYP
5	1239.8	80.7	1567	7 ACA61331	Aca61331 Human cyt
6	1239.8	80.7	1567	7 ADB25778	Adb25778 Human CYP
7	1238.2	80.6	1567	7 ACA61303	Aca61303 Human cyt
8	1234.6	80.3	1586	2 AAQ12893	Aaq12893 Debrisoqu
9	1225.6	79.7	1568	7 ACA61304	Aca61304 Human cyt
10	1211.2	78.8	1565	7 ACA61305	Aca61305 Human cyt
11	1201	78.1	1545	2 AAV19496	Aav19496 Cytochrom
12	1173.2	76.3	1494	2 AAQ87730	Aaq87730 Human aux
13	1173.2	76.3	1494	2 AAT28396	Aat28396 Human cyt
14	1173.2	76.3	1494	6 ABQ72216	Abq72216 Human CYP
15	1171.6	76.2	1494	2 AAQ87731	Aaq87731 Human aux
16	1171.6	76.2	1494	2 AAQ87729	Aaq87729 Human aux
17	1171.6	76.2	1494	2 AAT28395	Aat28395 Human cyt
18	1171.6	76.2	1494	2 AAT28397	Aat28397 Human cyt
19	1171.6	76.2	1494	2 AAT17388	Aat17388 Human der
20	1171.6	76.2	1494	8 ACF06056	Acf06056 Human cyt
21	1170	76.1	1494	2 AAQ87732	Aaq87732 Human aux
22	1170	76.1	1494	2 AAT28398	Aat28398 Human cyt
23	1103.8	71.8	1593	4 AAD09937	Aad09937 Human dru

24	828.6	53.9	1449	6 ABN97193	Abn97193 Gene #369
25	763	45.6	1270	6 ABN97194	Abn97194 Gene #369
26	730.6	47.5	2746	6 ABK63556	Abk63556 Rat seque
27	730.6	47.5	2746	9 ADB53689	Adb53689 Primary r
28	692.8	45.1	1630	9 ADB59020	Adb59020 Toxicity-
29	692.8	45.1	1630	9 ADB53778	Adb53778 Primary r
30	680	44.2	1698	9 ADB58102	Adb58102 Toxicity-
31	680	44.2	1698	9 ADB52582	Adb52582 Primary r
32	671	43.7	1628	9 ADB56180	Adb56180 Toxicity-
33	671	43.7	1628	9 ADB50703	Adb50703 Primary r
34	579.8	37.7	585	2 AAV19497	Aav19497 Cytochrom
35	561.2	36.5	603	2 AAQ12894	Aaq12894 Debrisoqu
36	551.8	35.9	591	4 AAH33792	Aah33792 Human col
37	506.4	32.9	1311	4 AAS59871	Aas59871 Human nov
38	368.6	24.0	701	6 ABT08965	Abt08965 Phase-1 R
39	331.4	21.6	10278	7 AAL53566	Aal53566 Genomic D
40	326.8	21.3	465	8 ACH15319	Ach15319 Human adu
41	321	20.9	477	2 AAV19498	Aav19498 Cytochrom
42	301.4	19.6	303	6 ABK93054	Abk93054 cDNA enco
43	290	18.9	4500	7 ADB25775	Adb25775 Human CYP
44	289.2	18.8	6472	6 ABQ72215	Abq72215 Human CYP
45	288.4	18.8	2170	7 ABT33958	Abt33958 Human pig

ALIGNMENTS

RESULT 1
AAL53565
ID AAL53565 standard; cDNA; 1537 BP.
XX
AC AAL53565;
XX
DT 30-JAN-2003 (first entry)
XX
DE cDNA encoding the drug-metabolising enzyme protein.
XX
KW Cytostatic; antiparkinsonian; gene therapy; cancer; Parkinson's disease;
KW human drug-metabolising protein; enzyme; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 78..1418
FT /*tag= a
FT /product= "Human drug-metabolising protein"
XX
PN WO200279233-A1.
XX
PD 10-OCT-2002.
XX
PF 01-APR-2002; 2002WO-US009738.
XX
PR 30-MAR-2001; 2001US-00820788.
XX
PA (PEKE) PE CORP NY.
PA (DFRA/) DI FRANCESCO V.
PA (BEAS/) BEASLEY E M.
XX
PI Shao W, Yan C;
XX
DR WPI; 2003-040649/03.
XX
DR P-PSDB; AAO26404.
XX
PT New human drug-metabolising proteins and nucleic acids related to the
PT Cytochrome P450 IID drug-metabolizing enzyme subfamily, useful for
PT treating a condition mediated by a human enzyme protein e.g., cancer.
XX
PS Claim 4; Fig 1; 72pp; English.
XX
CC The invention relates to a novel isolated polypeptide comprising a 446-
CC amino acid sequence or its allelic variant, orthologue or fragment. The
CC allelic variant or orthologue is encoded by a nucleic acid that

CC hybridises under stringent conditions to the opposite strand of the
 CC nucleic acid comprising a sequence having 1537 or 10278 bp. The fragment
 CC comprises at least 10 contiguous amino acids of the 446-amino acid
 CC sequence. The polypeptide is useful for preparing a pharmaceutical
 CC composition for treating a disease or condition mediated by a human
 CC enzyme protein, e.g. cancer or Parkinson's disease. This polynucleotide
 CC sequence represents the cDNA encoding the human drug-metabolising protein
 CC of the invention
 XX
 SQ Sequence 1537 BP; 301 A; 493 C; 443 G; 300 T; 0 U; 0 Other;

Query Match 100.0%; Score 1537; DB 7; Length 1537;
 Best Local Similarity 100.0%; Pred. No. 9.8e-283;
 Matches 1537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGCTGGTCTCTGTCCTGCTGGTGGGTGGGGTGCAGGTGTCTCAGAGGAGCCCAT 60
 DB 1 CCTGCTGGTCTCTGTCCTGCTGGTGGGTGGGGTGCAGGTGTCTCAGAGGAGCCCAT 60
 QY 61 TTGTAGTGGGAGGTATGGGCTAGAGCACTGGTCCCTGGCGGTGATGTGGCCA 120
 DB 61 TTGTAGTGGGAGGTATGGGCTAGAGCACTGGTCCCTGGCGGTGATGTGGCCA 120
 QY 121 TCTTCTGCTCTGCTGGTACCTGATGACCGCGCAACGCTGGGCTGCACGCTACTCAC 180
 DB 121 TCTTCTGCTCTGCTGGTACCTGATGACCGCGCAACGCTGGGCTGCACGCTACTCAC 180
 QY 181 CAGGCCCCCTGCACTGCGCGGCTGGGCAACCTGCTGCATGTGACATTCAGAACACAC 240
 DB 181 CAGGCCCCCTGCACTGCGCGGCTGGGCAACCTGCTGCATGTGACATTCAGAACACAC 240
 QY 241 CATACCTGCTCGACAGTTGGCGCCGCTTCGGGAGCGTGTTCAGCTGAGCTGGGCT 300
 DB 241 CATACCTGCTCGACAGTTGGCGCCGCTTCGGGAGCGTGTTCAGCTGAGCTGGGCT 300
 QY 301 GGACCCCGTGGTCTGCTCAATGGGCTGGCGGCGTGGGAGCGCTGTGTGACCCACG 360
 DB 301 GGACCCCGTGGTCTGCTCAATGGGCTGGCGGCGTGGGAGCGCTGTGTGACCCACG 360
 QY 361 GGGAGGACACCGCGGCGGCTGGCGGCGTGGGAGCGCTGTGTGACCCACG 420
 DB 361 GGGAGGACACCGCGGCGGCTGGCGGCGTGGGAGCGCTGTGTGACCCACG 420
 QY 421 GTTCCCAAGGAGCGCCCTTTTCGCCCAACGCTCTCTGAGCAAAAGCGTGAGCAACGTA 480
 DB 421 GTTCCCAAGGAGCGCCCTTTTCGCCCAACGCTCTCTGAGCAAAAGCGTGAGCAACGTA 480
 QY 481 TCGCTCTCCCTCACCTGCGGCGCGCTTCGAGTACGACGACCTCGCTTCCTCAGGCTGC 540
 DB 481 TCGCTCTCCCTCACCTGCGGCGCGCTTCGAGTACGACGACCTCGCTTCCTCAGGCTGC 540
 QY 541 TGGACCTAGCTCAGAGGAGCTGAAGAGGAGTGGGCTTTCTGCGGAGGTGCTGAATG 600
 DB 541 TGGACCTAGCTCAGAGGAGCTGAAGAGGAGTGGGCTTTCTGCGGAGGTGCTGAATG 600
 QY 601 CTGTCCCGCTCTCTGATATCCAGCGCTGGCTGGCAAGGCTCTAGCTTCGCAAGG 660
 DB 601 CTGTCCCGCTCTCTGATATCCAGCGCTGGCTGGCAAGGCTCTAGCTTCGCAAGG 660
 QY 661 CTTTCTGACCCAGCTGATGAGTCTAACTGAGCAGAGGATGACCTGGGACCCAGGCC 720
 DB 661 CTTTCTGACCCAGCTGATGAGTCTAACTGAGCAGAGGATGACCTGGGACCCAGGCC 720
 QY 721 AGCCCCCGGAGACTGAGTGGGCTTCCTGGCAGAGATGAGAGGAGCCCAAGGGGAAC 780
 DB 721 AGCCCCCGGAGACTGAGTGGGCTTCCTGGCAGAGATGAGAGGAGCCCAAGGGGAAC 780
 QY 781 CTGAGAGCAGCTTCAATGATGAGACCTGGCATAGTGGCTGACCTGCTCTGCGG 840
 DB 781 CTGAGAGCAGCTTCAATGATGAGACCTGGCATAGTGGCTGACCTGCTCTGCGG 840
 QY 841 GGATGGTGACCACTTCGACCCAGCTGGGCTGGGCGCTCTGCTCATGATCTCATCCGG 900
 DB 841 GGATGGTGACCACTTCGACCCAGCTGGGCTGGGCGCTCTGCTCATGATCTCATCCGG 900

DB 841 GGATGGTGACCACTTCGACCCAGCTGGGCGCTCTGCTCATGATCTCATCCGG 900
 QY 901 ATGTGACGCGCGTGTCCAAACAGGAGATCGACGACGTGATAGGCGAGTGGCGCACAG 960
 DB 901 ATGTGACGCGCGTGTCCAAACAGGAGATCGACGACGTGATAGGCGAGTGGCGCACAG 960
 QY 961 AGATGGGTGACCAAGCTCACATGCTTACACCACTGCGGTGATTCATGAGGTGAGCGCT 1020
 DB 961 AGATGGGTGACCAAGCTCACATGCTTACACCACTGCGGTGATTCATGAGGTGAGCGCT 1020
 QY 1021 TTGGGGACATGCTCCCTGGGTGTGACCCATATGACATCCCGTGAATCCGAAGTACAGG 1080
 DB 1021 TTGGGGACATGCTCCCTGGGTGTGACCCATATGACATCCCGTGAATCCGAAGTACAGG 1080
 QY 1081 GCTTCCGATCCCTAAAGGAGACGACACTCATCAACCACTGCTCATCGGTGCTGAAGGATG 1140
 DB 1081 GCTTCCGATCCCTAAAGGAGACGACACTCATCAACCACTGCTCATCGGTGCTGAAGGATG 1140
 QY 1141 AGGCGGTCTGGGAGAGCCCTTCCGCTTCCACCCGCAACACTTCTGATGCCAGGGCC 1200
 DB 1141 AGGCGGTCTGGGAGAGCCCTTCCGCTTCCACCCGCAACACTTCTGATGCCAGGGCC 1200
 QY 1201 ACTTTGTGAAGCGGAGCCCTTCTGCTTTCTCAGCAGGCGCGGTGATGCTCGGG 1260
 DB 1201 ACTTTGTGAAGCGGAGCCCTTCTGCTTTCTCAGCAGGCGCGGTGATGCTCGGG 1260
 QY 1261 AGCCCTGGCGCGCATGGAGCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320
 DB 1261 AGCCCTGGCGCGCATGGAGCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320
 QY 1321 TCTCGGTGCCACCTGAGCAGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
 DB 1321 TCTCGGTGCCACCTGAGCAGCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
 QY 1381 CCCCATCCCCCTATGAGCTTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 DB 1381 CCCCATCCCCCTATGAGCTTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 QY 1441 CTGCTCCCTAGCAGAGGCTCTAATGTACAAATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 DB 1441 CTGCTCCCTAGCAGAGGCTCTAATGTACAAATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 QY 1501 AA 1537
 DB 1501 AA 1537

RESULT 2
 ADB25779
 ID ADB25779 standard; DNA; 1567 BP.

XX ADB25779;
 AC ADB25779;
 XX 20-NOV-2003 (first entry)
 DT Human CYP2D6-related gene #4.
 DE human; mutant CYP2D6 gene; drug analysis; drug testing; gene; de.
 XX Homo sapiens.
 OS
 PH Key Location/Qualifiers
 FT CDS 1..1494
 FT /*tag= a
 FT /product= "Human CYP2D6-related protein #4"
 XX WO2003050282-A1.
 XX 19-JUN-2003.
 XX 05-DEC-2002; 2002WO-JP012748.
 XX 06-DEC-2001; 2001JP-00372548.

XX (TSUR) TSUMURA & CO.
XX Taniyama M, Ogawa K, Tsuchiya N, Hibino T;
XX WPI; 2003-505401/47.
XX Genetic polymorphisms of CYP2D6 gene in human population for analysis of
XX drug effect on individual patients and testing of new drugs.
XX Claim 8; Page 50-54; 75pp; Japanese.
XX The invention comprises mutant forms of the human CYP2D6 gene, containing
XX one or more of the following mutations G125A, C1858T, T2874C and C2875T.
XX The mutant human CYP2D6 genes of the invention are useful for analysing
XX the effect of drugs on individual patients and testing of new drugs. The
XX present DNA sequence represents a human gene of the invention.
XX SQ Sequence 1567 BP; 266 A; 526 C; 466 G; 309 T; 0 U; 0 Other;
Query Match 80.8%; Score 1241.4; DB 7; Length 1567;
Best Local Similarity 89.9%; Pred. No. 1.5e-226;
Matches 1408; Conservative 0; Mismatches 6; Indels 153; Gaps 1;
78 ATGGGGCTAGAGCACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCTGCTCTCTGGTG 137
1 ATGGGGCTAGAGCACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCTGCTCTCTGGTG 60
138 GACCTGATGACACCGGGCCCAACGCTGGGTGACAGCTACTACAGGCCCTGCGCACTG 197
61 GACCTGATGACACCGGGCCCAACGCTGGGTGACAGCTACTACAGGCCCTGCGCACTG 120
198 CCGGGCTGGGCAACCTGCTGATGTGGAATTCAGAAACACCATATCTGCTTCGACCAG 257
121 CCGGGCTGGGCAACCTGCTGATGTGGAATTCAGAAACACCATATCTGCTTCGACCAG 180
258 TTGGCGGCGCGCTTGGGGAACGTGTTAGCTGAGCTGGCTGGAACGCGGTGCTG 317
181 TTGGCGGCGCGCTTGGGGAACGTGTTAGCTGAGCTGGCTGGAACGCGGTGCTG 240
318 CTCAATGGCTGGCGCGCTGGCGGAGCGCTGGTACCCAGCGGAGGACACCGCCGAC 377
241 CTCAATGGCTGGCGCGCTGGCGGAGCGCTGGTACCCAGCGGAGGACACCGCCGAC 300
378 CGCCCGCTGTCCTCATACCCAGATCTCTGGGTGTTGGGCGCGCTTCCCAA 428
301 CGCCCGCTGTCCTCATACCCAGATCTCTGGGTGTTGGGCGCGCTTCCCAA 360
429 428
361 CTGGCGCGCTATGGGCGCGCTGGCGGAGCAGAGCGCTTCTCGTGTCCACCTTGGCG 420
429 428
421 AACTTGGGCTGGGCAAGATCGCTGGAGCAGTGGTGAACGAGGAGGCGCGCTT 480
429 428
481 TGTGGCGCTTGGCAACCACTCCGAGCGCCCTTTCGCCCAACGCTCTTGGACAAA 540
465 GCCGTGAGCAACGTGATCGCTTCCCTCACTCGCGGCGCGCTTCGAGTACGACACCT 524
541 GCCGTGAGCAACGTGATCGCTTCCCTCACTCGCGGCGCGCTTCGAGTACGACACCT 600
525 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGACTGAAGGAGTCTGGGCTTCTG 584
601 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGACTGAAGGAGTCTGGGCTTCTG 660
585 CGCGAGGTGCTGAATGCTGCTCCCGCTCTCTGCTGATATCCAGAGCGCTGCTGGCAAGTTC 644
661 CGCGAGGTGCTGAATGCTGCTCCCGCTCTCTGCTGATATCCAGAGCGCTGCTGGCAAGTTC 720
645 CTACGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGCTGCTAACTGAGCACAGGATG 704

Db 721 CTACGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGCTGCTAACTGAGCACAGGATG 780
Qy 705 ACTTGGGACCCAGCCAGCCCGAGACCTGACTGAGGCTTCTCTGCGACAGATGGAG 764
Db 781 ACTTGGGACCCAGCCAGCCCGAGACCTGACTGAGGCTTCTCTGCGACAGATGGAG 840
Qy 765 AAGCCAAAGGGAAACCTCTGAGAGCAGCTTCAATGATGAGAACTTGCATAGTGGTCT 824
Db 841 AAGCCAAAGGGAAACCTCTGAGAGCAGCTTCAATGATGAGAACTTGCATAGTGGTCT 900
Qy 825 GACCTGTTCTCTGCGGGGATGTTACCACTCTGACCACTGCTGGCTTGGGCTCTCTGCTC 884
Db 901 GACCTGTTCTCTGCGGGGATGTTACCACTCTGACCACTGCTGGCTTGGGCTCTCTGCTC 960
Qy 885 ATGATCTTACATCTCGGATGTCAGCGCGCTGTCCAAACAGGAGATCGACGACCTGATAGG 944
Db 961 ATGATCTTACATCTCGGATGTCAGCGCGCTGTCCAAACAGGAGATCGACGACCTGATAGG 1020
Qy 945 CAGGTGCGGGCCAGCAGAGATGGTGAACAGGCTCAGATGCCCTACACCACTGCCGTGATT 1004
Db 1021 CAGGTGCGGGCCAGCAGAGATGGTGAACAGGCTCAGATGCCCTACACCACTGCCGTGATT 1080
Qy 1005 CATGAGTGCAGCGCTTTTGGGGAACATGCTCCCTGGGTGTGACCCATATGACATCCCT 1064
Db 1081 CATGAGTGCAGCGCTTTTGGGGAACATGCTCCCTGGGTGTGACCCATATGACATCCCT 1140
Qy 1065 GACATCGAAGTACAGGGCTTCGGCATCTTAAAGGAAACGACACTCATCAACCACTGTCA 1124
Db 1141 GACATCGAAGTACAGGGCTTCGGCATCTTAAAGGAAACGACACTCATCAACCACTGTCA 1200
Qy 1125 TCGGTGCTGAAGGATGAGGCGCTCTGGGAGAGCGCTTCCGCTTCCACCCGAACTTC 1184
Db 1201 TCGGTGCTGAAGGATGAGGCGCTCTGGGAGAGCGCTTCCGCTTCCACCCGAACTTC 1260
Qy 1185 CTGATGCCAGGGCCACTTTGTGAAGCGGAGGCGCTTCTGCTTCTCAGAGGCGCG 1244
Db 1261 CTGATGCCAGGGCCACTTTGTGAAGCGGAGGCGCTTCTGCTTCTCAGAGGCGCG 1320
Qy 1245 CTGATGCCCTCGGGGAGCGCTTGGCGGAGCGCTTCCGCTTCTTCACTCCCTCTG 1304
Db 1321 CTGATGCCCTCGGGGAGCGCTTGGCGGAGCGCTTCCGCTTCTTCACTCCCTCTG 1380
Qy 1305 CTGAGCACTTTCAGCTTCTCGGTGCCCTCTGAGACCGCGCCGAGCCAGCCATGGTGC 1364
Db 1381 CTGAGCACTTTCAGCTTCTCGGTGCCCTCTGAGACCGCGCCGAGCCAGCCATGGTGC 1440
Qy 1365 TTTGCTTCTCTGTTGACCCCATCCCTTATGAGCTTTGTGCTGTGCCCGCTAGAAATGG 1424
Db 1441 TTTGCTTCTCTGTTGAGCCCATCCCTTATGAGCTTTGTGCTGTGCCCGCTAGAAATGG 1500
Qy 1425 GTACCTAGTCCCGAGCTGCTCCCTAGCCAGAGGCTCTAATGTACAATAAGCAATGG 1484
Db 1501 GTACCTAGTCCCGAGCTGCTCCCTAGCCAGAGGCTCTAATGTACAATAAGCAATGG 1560
Qy 1485 TAGTTCC 1491
Db 1561 TAGTTCC 1567

RESULT 3

ADB25777

ID ADB25777 standard; DNA; 1567 BP.

XX ADB25777;

AC ADB25777;
XX 20-NOV-2003 (first entry)

DE Human CYP2D6-related gene #2.

XX human; mutant CYP2D6 gene; drug analysis; drug testing; gene; db.
XX Homo sapiens.

XX Key Location/Qualifiers
FH 1..1494
FT /*tag= a
FT /product= "Human CYP2D6-related protein #2"
XX WO2003050282-A1.
XX PD 19-JUN-2003.
XX PF 05-DEC-2002; 2002WO-JP012748.
XX PR 06-DEC-2001; 2001JP-00372548.
XX PA (TSUR) TSUMURA & CO.
XX Taniyama M, Ogawa K, Tsuchiya N, Hibino T;
XX WPI; 2003-505401/47.
XX P-PSDB; ADB25832.
XX Genetic polymorphisms of CYP2D6 gene in human population for analysis of
XX drug effect on individual patients and testing of new drugs.
XX PS 'Claim 8; Page 43-46; 75pp; Japanese.
XX CC The invention comprises mutant forms of the human CYP2D6 gene, containing
XX one or more of the following mutations G125A, C1858T, T2874C and C2875T.
XX The mutant human CYP2D6 genes of the invention are useful for analysing
XX the effect of drugs on individual patients and testing of new drugs. The
XX present DNA sequence represents a human gene of the invention.
XX SQ Sequence 1567 BP; 266 A; 527 C; 464 G; 310 T; 0 U; 0 Other;

Query Match 80.8%; Score 1241.4; DB 7; Length 1567;
Best Local Similarity 89.9%; Pred. No. 1.5e-226;
Matches 1408; Conservative 0; Mismatches 6; Indels 153; Gaps 1;

QY 78 ATGGGGCTAGAGCACTGGTGGCCCTGGCGTGATAGTGGCCATCTTCTGCTCTGCTG 137
DB |||||
QY 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCGTGATAGTGGCCATCTTCTGCTCTGCTG 60
DB |||||
QY 138 GACCTGATGACACGGCGCAACGCTGGCTGACGCTACTACACAGGCCCTGCACTG 197
DB |||||
QY 61 GACCTGATGACACGGCGCAACGCTGGCTGACGCTACTACACAGGCCCTGCACTG 120
DB |||||
QY 198 CCCGGCTGGGCAACCTGCTGATGGAATTCAGAAACACACCACTACTGCTTCGACCA 257
DB |||||
QY 121 CCCGGCTGGGCAACCTGCTGATGGAATTCAGAAACACACCACTACTGCTTCGACCA 180
DB |||||
QY 258 TTGGCGCGCGCTTCGGGGACGTGTTACGCTGACGCTGGCTGACCGCGGTGCTG 317
DB |||||
QY 181 TTGGCGCGCGCTTCGGGGACGTGTTACGCTGACGCTGGCTGACCGCGGTGCTG 240
DB |||||
QY 318 CTCATAGGCTGGCGCGCTGGCGGACGCTGGTACCCAGCGGAGGACACCGCCGAC 377
DB |||||
QY 241 CTCATAGGCTGGCGCGCTGGCGGACGCTGGTACCCAGCGGAGGACACCGCCGAC 300
DB |||||
QY 378 CGCCGCGCTGTGCCCATCACCGATCTCTGGGTTTGGGCGCGCTTCCTCCAA- 428
DB |||||
QY 301 CGCCGCGCTGTGCCCATCACCGATCTCTGGGTTTGGGCGCGCTTCCTCCAAAGGGGTG 360
DB |||||
QY 429 ----- 428
DB 361 CTGGCGCGCTATGGGCGCGCTGGCGGACGAGGCGCTTCTCGCTCTCCACTTGGCG 420
QY 429 ----- 428
DB 421 AACTTGGGCTGGGCAAGAGTCGCTGGAGCAGTGGGTGACCGAGGAGCGCGCTGCTT 480
QY 429 -----GGAGCGCGCTTTTCGCGCGCGCGCTCTCTTGGACAAA 464
DB |||||
QY 481 TGTGCGCGCTTTCGCGCAACCACTTCGCGACGCGCGCTTTCGCGCGCGCTCTCTTGGACAAA 540
DB |||||

QY 465 GCCGTGAGCAAGCTGATCGCCTCCCTCACCTGCGGGCGCGCTTTCGAGTACGACGACCT 524
DB |||||
QY 541 GCCGTGAGCAAGCTGATCGCCTCCCTCACCTGCGGGCGCGCTTTCGAGTACGACGACCT 600
DB |||||
QY 525 CGCTTCTCTCAGGCTGCTGGACCTAGCTCAGGAGGAGCTGAAGAGGAGGAGTCCGGCTTCTG 584
DB |||||
QY 601 CGCTTCTCTCAGGCTGCTGGACCTAGCTCAGGAGGAGCTGAAGAGGAGGAGTCCGGCTTCTG 660
DB |||||
QY 585 CGCGAGGTGCTGAATGCTGCTCCCGTCTCTCTGATATCCAGCGCTGGCTGGCAAGGTC 644
DB |||||
QY 661 CGCGAGGTGCTGAATGCTGCTCCCGTCTCTCTGATATCCAGCGCTGGCTGGCAAGGTC 720
DB |||||
QY 645 CTACGCTTCCAAAGGCTTCTCTGACCCAGCTGATGAGCTGCTAACTGACACAGGATG 704
DB |||||
QY 721 CTACGCTTCCAAAGGCTTCTCTGACCCAGCTGATGAGCTGCTAACTGACACAGGATG 780
DB |||||
QY 705 ACCTGGGACCCAGCGCCCGAGACCTGCTGAGGCTTCTCTGGCAGAGATGGAG 764
DB |||||
QY 781 ACCTGGGACCCAGCGCCCGAGACCTGCTGAGGCTTCTCTGGCAGAGATGGAG 840
DB |||||
QY 765 AAGCCCAAGGGGAACCTGAGAGGAGCTTCAATGATGAGAACTTCGATATGCTGCT 824
DB |||||
QY 841 AAGCCCAAGGGGAACCTGAGAGGAGCTTCAATGATGAGAACTTCGATATGCTGCT 900
DB |||||
QY 825 GACCTGTTCTCTGCGGGATGGTGACCACTCGACCTGCGCTGGCTGGGGCTCTCTGCTC 884
DB |||||
QY 901 GACCTGTTCTCTGCGGGATGGTGACCACTCGACCTGCGCTGGCTGGGGCTCTCTGCTC 960
DB |||||
QY 885 ATGATCTCTACATCCGAGTGTGACGCGCGTGTCAACAGAGGAGATCGACGCTGATAGG 944
DB |||||
QY 961 ATGATCTCTACATCCGAGTGTGACGCGCGTGTCAACAGAGGAGATCGACGCTGATAGG 1020
DB |||||
QY 945 CAGTGGCGCGACACAGAGATGGGTGACAGGCTCACATGCCCTACACACTGCCGTGATT 1004
DB |||||
QY 1021 CAGTGGCGCGACACAGAGATGGGTGACAGGCTCACATGCCCTACACACTGCCGTGATT 1080
DB |||||
QY 1005 CATCAGGTGACGCGCTTGGGGACATGTCCTCCCTGGGTGACCCATATGACATCCCGT 1064
DB |||||
QY 1081 CATCAGGTGACGCGCTTGGGGACATGTCCTCCCTGGGTGACCCATATGACATCCCGT 1140
DB |||||
QY 1065 GACATCGAAGTACAGGGCTTCGCGATCCCTAAGGGAACGACATCATCACCAACTGTCA 1124
DB |||||
QY 1141 GACATCGAAGTACAGGGCTTCGCGATCCCTAAGGGAACGACATCATCACCAACTGTCA 1200
DB |||||
QY 1125 TCGGTGCTGAAGGATGAGGCGCTCTGGGAGAGCCCTTCGCTTCCACCCCGAACACTTC 1184
DB |||||
QY 1201 TCGGTGCTGAAGGATGAGGCGCTCTGGGAGAGCCCTTCGCTTCCACCCCGAACACTTC 1260
DB |||||
QY 1185 CTGGATGCCAGGGCCACTTTGTGAAGCGGAGGCTTCTCTGCTTCTCAGCAGGCGCG 1244
DB |||||
QY 1261 CTGGATGCCAGGGCCACTTTGTGAAGCGGAGGCTTCTCTGCTTCTCAGCAGGCGCG 1320
DB |||||
QY 1245 CGTGCATGCTCTGGGAGAGCCCTTCGCGCGATGAGGCTTCTCTTCTTCTTCACTTCCCTG 1304
DB |||||
QY 1321 CGTGCATGCTCTGGGAGAGCCCTTCGCGCGATGAGGCTTCTCTTCTTCTTCACTTCCCTG 1380
DB |||||
QY 1305 CTGACGACTTTCAGGCTCTCGGTGCCACTGGAAGCGCCCGCGCGAGCAGCAGTGGTCTC 1364
DB |||||
QY 1381 CTGACGACTTTCAGGCTCTCGGTGCCACTGGAAGCGCCCGCGCGAGCAGCAGTGGTCTC 1440
DB |||||
QY 1365 TTTGCTTCTTCGCTGAGCCCATCCCTTATCAGCTTGTGTGTGCTGCTGCTGCTGCTGCTG 1424
DB |||||
QY 1441 TTTGCTTCTTCGCTGAGCCCATCCCTTATCAGCTTGTGTGTGCTGCTGCTGCTGCTGCTG 1500
DB |||||
QY 1425 GTACTAGTCTCCCGAGCTGCTCCCTTAGCCAGAGGCTCTAATGTACATAAAGCAATGTGG 1484
DB |||||
QY 1501 GTACTAGTCTCCCGAGCTGCTCCCTTAGCCAGAGGCTCTAATGTACATAAAGCAATGTGG 1560
DB |||||
QY 1485 TAGTTCC 1491
DB |||||
QY 1561 TAGTTCC 1567

```
RESULT 4
ADB25776 standard; DNA; 1567 BP.
XX AC ADB25776;
XX DT 20-NOV-2003 (first entry)
XX DE Human CYP2D6-related gene #1.
XX KW human; mutant CYP2D6 gene; drug analysis; drug testing; gene; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT 1. .1494
XX FT /*tag= a
XX FT /product= "Human CYP2D6-related protein #1"
XX PN WO2003050282-A1.
XX PD 19-JUN-2003.
XX PF 05-DEC-2002; 2002WO-JP012748.
XX PR 06-DEC-2001; 2001JP-00372548.
XX PS (TSUR ) TSUMURA & CO.
XX PI Taniyama M, Ogawa K, Tsuchiya N, Hibino T;
XX WPI; 2003-505401/47.
XX DR P-PSDB; ADB25831.
XX CC The invention comprises mutant forms of the human CYP2D6 gene, containing
XX one or more of the following mutations G125A, C1858T, T2874C and C2875T.
XX CC The mutant human CYP2D6 genes of the invention are useful for analysing
XX the effect of drugs on individual patients and testing of new drugs. The
XX present DNA sequence represents a human gene of the invention.
XX SQ Sequence 1567 BP; 266 A; 527 C; 464 G; 310 T; 0 U; 0 Other;

Query Match 80.8%; Score 1241.4; DB 7; Length 1567;
Best Local Similarity 89.9%; Pred. No. 1.5e-226;
Matches 1408; Conservative 0; Mismatches 6; Indels 153; Gaps 1;

QY 78 ATGGGGCTAGAGCACTGGTGGCCCTGGCCGTGATAGTGGCCATCTTCTGCTCTGTG 137
DB 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCCGTGATAGTGGCCATCTTCTGCTCTGTG 60
QY 138 GACCTGATCACCGCGGCAACCGCTGGGCTGCAGCTACTCACAGGCCCTGCGCACTG 197
DB 61 GACCTGATCACCGCGGCAACCGCTGGGCTGCAGCTACTCACAGGCCCTGCGCACTG 120
QY 198 CCGGGCTGGGCAACCTGCTGATGTGGACTTCCAGAACACACCATACTGCTTCGACCAG 257
DB 121 CCGGGCTGGGCAACCTGCTGATGTGGACTTCCAGAACACACCATACTGCTTCGACCAG 180
QY 258 TTGCGGCGCGCTTCGGGAGCGTGTTCAGCTGCGAGCTGGCTGAGCGCGGTGTGCTG 317
DB 181 TTGCGGCGCGCTTCGGGAGCGTGTTCAGCTGCGAGCTGGCTGAGCGCGGTGTGCTG 240
QY 318 CTCATATGGGCTGGCGGCGCTGGCGGCGCTGGTGAACCCAGCGGAGGACACCGCGAC 377
DB 241 CTCATATGGGCTGGCGGCGCTGGCGGCGCTGGTGAACCCAGCGGAGGACACCGCGAC 300
QY 378 CGCCCGCCTGTGCCCATCACCCAGATCTCTGGGTTTTTGGGCCGCGTTCCCAA----- 428

301 CGCCCGCCTGTGCCCATCACCCAGATCTGGGTTTTTGGGCCGCGTTCCCAAAGGGTGTTC 360
429 ----- 428
361 CTGGCGCGCTATGGCCCGCGTGGCGCGAGCAGAGGCGCTTCTCCGTCTCCACCTTGGCG 420
429 ----- 428
421 AACTTGGGCGCTGGGCAAGAGTCGCTGGAGCAGTGGGTGACCGAGGAGGCGCGCTTTCG 480
429 ----- 428
481 TGTGCGCGCTTTCGCCAACCACTCCGGGACGCGCTTTCGCCCGCAACGCTCTCTTGACA 540
465 GCGGTGAGCAACAGTGTGCTGCTGCGGCGCGCTTTCGAGGAGTACGAGCAGCCTT 524
541 GCGGTGAGCAACAGTGTGCTGCTGCGGCGCGCTTTCGAGGAGTACGAGCAGCCTT 600
525 CGCTTCTCAGGCTGCTGGACCTAGCTCAGGAGGAGTGAAGGAGGAGTTCGGGCTTTCG 584
601 CGCTTCTCAGGCTGCTGGACCTAGCTCAGGAGGAGTGAAGGAGGAGTTCGGGCTTTCG 660
585 GCGGAGGTGCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 644
661 GCGGAGGTGCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
645 CTACGCTTCCAAAGGCTTCTGCTGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 704
721 CTACGCTTCCAAAGGCTTCTGCTGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
705 ACCTGGGACCCAGCCCGCCCGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 764
781 ACCTGGGACCCAGCCCGCCCGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
765 AAGGCCAAGGGGAACTCTGAGAGCAGCTTCAATGATGAGAACTCTGCGCATAGTGTGGCT 824
841 AAGGCCAAGGGGAACTCTGAGAGCAGCTTCAATGATGAGAACTCTGCTGATAGTGTGGCT 900
825 GACCTGTTCTCTGCGGGATGCTGACCACTGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTG 884
901 GACCTGTTCTCTGCGGGATGCTGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
885 ATGATCTCTACATCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
961 ATGATCTCTACATCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
945 CAGTGTGCGGCAACAGAGATGGGTGACCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1004
1021 CAGTGTGCGGCAACAGAGATGGGTGACCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
1005 CATGAGGTGACAGCGCTTTCGGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1064
1081 CATGAGGTGACAGCGCTTTCGGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
1065 GACATCGAAGTACAGGCTTTCGGATCCCTAAGGGAACGACACTCATCACCAACCTGTCA 1124
1141 GACATCGAAGTACAGGCTTTCGGATCCCTAAGGGAACGACACTCATCACCAACCTGTCA 1200
1125 TCGGTGCTGAAGGATGAGGCGCTCTGGGAGAGCGCTTTCGGCTTTCACCCCGAACAACCTTC 1184
1201 TCGGTGCTGAAGGATGAGGCGCTCTGGGAGAGCGCTTTCGGCTTTCACCCCGAACAACCTTC 1260
1185 CTGATGCCAGGGGCACTTTCGAGCGGAGGCGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1244
1261 CTGATGCCAGGGGCACTTTCGAGCGGAGGCGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
1245 CGTGATGCCCTCGGGAGCGCTTTCGCGGAGGCGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1304
1321 CGTGATGCCCTCGGGAGCGCTTTCGCGGAGGCGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
1305 CTGAGCACTTCAGCTTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1364
```

Db 1381 CTGACGACTTCAGCTTCTCGTGCCCACTGACAGCCCGCCGACCCATGATGGTGC 1440
Qy 1365 TTTGCTTTCTGCTGACCCCATCCCTCATGAGCTTTGCTGTCGCCCGCTAGATGGG 1424
Db 1441 TTTGCTTTCTGCTGACCCCATCCCTCATGAGCTTTGCTGTCGCCCGCTAGATGGG 1500
Qy 1425 GTACCTAGTCCCGACGCTGCTCCCTAGCAGAGGCTCTAAATGACAAATAAGCAATGTGG 1484
Db 1501 GTACCTAGTCCCGACGCTGCTCCCTAGCAGAGGCTCTAAATGACAAATAAGCAATGTGG 1560
Qy 1485 TAGTTC 1491
Db 1561 TAGTTC 1567

RESULT 5

ID ACA61331 standard; cDNA; 1567 BP.

XX ACA61331;

AC ACA61331;

XX 16-JUL-2003 (first entry)

XX Human cytochrome p450 gene CYP2D6 G5799C variant, cDNA.

XX Human; ss; gene; cytochrome P450; CYP2D6; chromosome 22; SNP;

XX single nucleotide polymorphism; drug metabolism; cardiovascular disorder;

XX psychiatric disorder; drug sensitivity.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1494

XX /tag= a

XX /product= "CYP2D6 protein"

XX replace(1457,G)

XX /tag= b

XX /standard_name= "Single nucleotide polymorphism"

XX EP1281755-A2.

XX 05-FEB-2003.

XX 16-JUL-2002; 2002EP-00254972.

XX 31-JUL-2001; 2001US-0309111P.

XX (PF12) PFIZER PROD INC.

XX Milos PM, Webb SM;

XX WPI; 2003-373769/36.

XX P-PSDB; AB009598.

XX New cytochrome P450 2D6 gene variants and polypeptides, useful for

XX determining if a subject has or is at risk of developing a drug

XX sensitivity condition or disorder that is associated with an aberrant

XX CYP2D6 activity.

XX Disclosure; Fig 7; 88pp; English.

XX The invention relates to an isolated nucleic acid comprising a cytochrome

XX P450 2D6 gene variant, e.g. G5799C or C5816AT (referring to the genomic

XX sequence) or the same variant nucleotide in the corresponding cDNA

XX sequences). Also included are probes, primers (allele specific

XX oligonucleotides) and arrays used to detect and/or amplify the CYP2D6

XX gene polymorphic regions, the variant polypeptides, antibodies which are

XX capable of distinguishing between the variant and wild-type polypeptides,

XX determining whether a subject has a genetic deficiency for metabolising a

XX drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and

XX determining whether an individual is susceptible to being a poor

XX metaboliser of drugs. The DNA probe is useful for hybridising to a

XX variant form of the CYP2D6 gene. The primer is useful for amplifying the

CC C5816TA allelic variant. The allele specific nucleotide is useful for the
CC detection of the C5816TA allelic variant. The methods are useful for
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and
CC determining if an individual is susceptible to being a poor metaboliser
CC of drugs. The nucleic acids are useful as probes or primers for
CC determining whether a subject has a genetic deficiency for metabolising
CC drugs that are substrates of P450 CYP2D6. The methods are useful for
CC determining if a subject has or is at risk of developing a drug
CC sensitivity condition or disorder that is associated with an aberrant
CC CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
CC aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
CC appropriate drugs or determining the course of treatment to administer to
CC a subject to treat cardiovascular or psychiatric disorders, or for
CC treating a subject with a drug sensitivity or disorder associated with a
CC specific allelic variant of a polymorphic region of the CYP2D6 gene. The
CC antibodies are useful for monitoring CYP2D6 protein levels in an
CC individual for determining whether a subject has a disease or conditions
CC associated with an aberrant CYP2D6 protein level. The gene is located on
CC human chromosome 22. The present sequence is the CYP2D6 cDNA carrying the
CC G5799C/C5816TA double variation
XX

SQ Sequence 1567 BP; 267 A; 528 C; 464 G; 308 T; 0 U; 0 Other;

Query Match 80.7%; Score 1239.8; DB 7; Length 1567;

Best Local Similarity 89.8%; Pred. No. 3e-226;

Matches 1407; Conservative 0; Mismatches 7; Indels 153; Gaps 1;

Qy 78 ATGGGGCTAGAACACTGGTGGCCCTGGCCGTGATAGTGGCCATCTTCCTGCTCTGGTG 137

Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCCGTGATAGTGGCCATCTTCCTGCTCTGGTG 60

Qy 138 GACCTGATGACCGCGGCCAACGCTGGCTGCACTCTCACCAGGCCCCCTGCCACTG 197

Db 61 GACCTGATGACCGCGGCCAACGCTGGCTGCACTCTCACCAGGCCCCCTGCCACTG 120

Qy 198 CCCGGGCTGGGCAACCTGCTGCACTGTCAGAACACACCATATCTGTTGACACAG 257

Db 121 CCCGGGCTGGGCAACCTGCTGCACTGTCAGAACACACCATATCTGTTGACACAG 180

Qy 258 TTGCGGCGCGCTTTCGGGGACGTTTCAGCTGTCAGCTGCGAGCCCGGTGGTGGTG 317

Db 181 TTGCGGCGCGCTTTCGGGGACGTTTCAGCTGTCAGCTGCGAGCCCGGTGGTGGTG 240

Qy 318 CTCATGCGCTGGCGCGCGCTGCGAGCGCTGTCACCCACCGGAGGACACCGCCGAC 377

Db 241 CTCATGCGCTGGCGCGCGCTGCGAGCGCTGTCACCCACCGGAGGACACCGCCGAC 300

Qy 378 CGCCCGCTGTGCCCATCACTGAGATCTGGGTTTTCGGGCGCGCTTCCCAA----- 428

Db 301 CGCCCGCTGTGCCCATCACTGAGATCTGGGTTTTCGGGCGCGCTTCCCAAAGGGTGTTC 360

Qy 429 ----- 428

Db 361 CTGCGCGCTATGGCGCGCGTGGCGGAGCAGAGGGCGTTCTCCGTGTCCACCTTGGCG 420

Qy 429 ----- 428

Db 421 AACTTGGGCTGGGCAAGAAAGTTCGCTGGAGCAGTGGGTGACCCGAGGAGCGCGCTGCTT 480

Qy 429 -----GGACGCGCGCTTTCGCCCCAACCGGTCTCTTGGACAAA 464

Db 481 TGTGCGCGCTTCGGCAACCACTTCGCGAGCGCGCTTTCGCCCCAACCGGTCTCTTGGACAAA 540

Qy 465 GCCGTGAGCAACGATGATGCGCTTCCCTCACTGCGGGCGCGCTTTCGAGTACGACGACCT 524

Db 541 GCCGTGAGCAACGATGATGCGCTTCCCTCACTGCGGGCGCGCTTTCGAGTACGACGACCT 600

Qy 525 CGCTTCTCAGGCTGCTGGACCTAGCTCAGAGGAGTGAAGGAGGAGTGGGCTTCTG 584

Db 601 CGCTTCTCAGGCTGCTGGACCTAGCTCAGAGGAGTGAAGGAGGAGTGGGCTTCTG 660

Qy 585 CGCAGGCTGTGAATGTCTCTCCCGCTCTCTGTCATATCCAGCGCTGGCTGGCAAGGTC 644

Db 661 CGGAGGTGCTGAATGCTGTCCTCCGCTCTCTGATATCCAGCGTGGCTGCAAGGTC 720
 Qy 645 CTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGCTGCTAACTGAGCACAGATG 704
 Db 721 CTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGCTGCTAACTGAGCACAGATG 780
 Qy 705 ACCTGGACCCAGCCAGCCGCGGAGACTGACTGAGGCTTCTCTGGCAGAGATGGAG 764
 Db 781 ACCTGGACCCAGCCAGCCGCGGAGACTGACTGAGGCTTCTCTGGCAGAGATGGAG 840
 Qy 765 AAGGCCAAGGGGAACTGAGAGAGCTTCAATGATGAGAACTGCGCATAGTGGTGGCT 824
 Db 841 AAGGCCAAGGGGAACTGAGAGAGCTTCAATGATGAGAACTGCGCATAGTGGTGGCT 900
 Qy 825 GACCTGTTCTCTGCGGGATGGTGACCACTGACCACTGCGGCTGGGCTTCTGCTC 884
 Db 901 GACCTGTTCTCTGCGGGATGGTGACCACTGACCACTGCGGCTGGGCTTCTGCTC 960
 Qy 885 ATGATCTATATCCGGATGTCAGGCGGCTGTCACAGGAGATCGACGCTGATAGGG 944
 Db 961 ATGATCTATATCCGGATGTCAGGCGGCTGTCACAGGAGATCGACGCTGATAGGG 1020
 Qy 945 CAGGTGCGGACACAGAGATGGTGACCACTGACCACTGCGGCTGGGCTTCTGCTC 1004
 Db 1021 CAGGTGCGGACACAGAGATGGTGACCACTGACCACTGCGGCTGGGCTTCTGCTC 1080
 Qy 1005 CATGAGGTGAGGCTTTTGGGACATGCTCCCTGCTGCTGACCCATATGATCCCT 1064
 Db 1081 CATGAGGTGAGGCTTTTGGGACATGCTCCCTGCTGCTGACCCATATGATCCCT 1140
 Qy 1065 GACATCAAGTACAGGCTTTCCGATCCCTTAAGGGAACGACATCATACCACTGTCA 1124
 Db 1141 GACATCAAGTACAGGCTTTCCGATCCCTTAAGGGAACGACATCATACCACTGTCA 1200
 Qy 1125 TCGGTGCTGAAGATGAGGCGCTTGGGAGAGCCCTTCCGCTTCCACCCGACACTTC 1184
 Db 1201 TCGGTGCTGAAGATGAGGCGCTTGGGAGAGCCCTTCCGCTTCCACCCGACACTTC 1260
 Qy 1185 CTGATGCTCCAGGCGCACTTTGTGAAGCGGAGGCTTCTGCTTTCTCAGCAGGCGC 1244
 Db 1261 CTGATGCTCCAGGCGCACTTTGTGAAGCGGAGGCTTCTGCTTTCTCAGCAGGCGC 1320
 Qy 1245 CGTGATGCTCGGAGGCGCTTGGCGGATGAGGCTTCTTCTTCTTCACTCCCTG 1304
 Db 1321 CGTGATGCTCGGAGGCGCTTGGCGGATGAGGCTTCTTCTTCTTCACTCCCTG 1380
 Qy 1305 CTGACGACTTACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1364
 Db 1381 CTGACGACTTACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 Qy 1365 TTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1424
 Db 1441 TTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 Qy 1425 GTACCTAGTCCCGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1484
 Db 1501 GTACCTAGTCCCGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 Qy 1485 TAGTTCC 1491
 Db 1561 TAGTTCC 1567

RESULT 6
 ADB25778
 ID ADB25778 standard; DNA; 1567 BP.
 XX
 AC ADB25778;
 XX
 DT 20-NOV-2003 (first entry)
 XX Human CYP2D6-related gene #3.
 DE

human; mutant CYP2D6 gene; drug analysis; drug testing; gene; ds.
 XX
 KW Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT 1.1494
 FT /tag= a
 FT /product= "Human CYP2D6-related protein #3"
 XX
 PN WO2003050282-A1.
 XX
 PD 19-JUN-2003.
 XX
 PF 05-DEC-2002; 2002WO-JP012748.
 XX
 PR 06-DEC-2001; 2001JP-00372548.
 XX
 PA (TSUR) TSUMURA & CO.
 XX
 PI Taniyama M, Ogawa K, Tauchiya N, Hibino T;
 XX
 DR WPI: 2003-505401/47.
 XX
 DR P-PSDB; AD25833.
 XX
 PT Genetic polymorphisms of CYP2D6 gene in human population for analysis of
 PT drug effect on individual patients and testing of new drugs.
 XX
 PS Claim 8; Page 46-50; 75pp; Japanese.
 XX
 CC The invention comprises mutant forms of the human CYP2D6 gene, containing
 CC one or more of the following mutations G125A, C1858T, T2874C and C2875T.
 CC The mutant human CYP2D6 genes of the invention are useful for analysing
 CC the effect of drugs on individual patients and testing of new drugs. The
 CC present DNA sequence represents a human gene of the invention.
 XX
 SQ Sequence 1567 BP; 266 A; 528 C; 465 G; 308 T; 0 U; 0 Other;
 XX
 Query Match 80.7%; Score 1239.8; DB 7; Length 1567;
 Best Local Similarity 89.8%; Pred. No. 3e-226;
 Matches 1407; Conservative 0; Mismatches 7; Indels 153; Gaps 1;
 QY 78 ATGGGGCTAGAGACACTGGTCCCTGGCGTGTAGTGGCCATCTTCTGCTCTGCTG 137
 DB 1 ATGGGGCTAGAGACACTGGTCCCTGGCGTGTAGTGGCCATCTTCTGCTCTGCTG 60
 QY 138 GACCTGATGACACGGGCGCAACGCTGGGCTGACGCTACTACACGAGCCCTTGCAC 197
 DB 61 GACCTGATGACACGGGCGCAACGCTGGGCTGACGCTACTACACGAGCCCTTGCAC 120
 QY 198 CCGGGCTGGGCAACCTGCTGATGCTGACTTCCAGAACACACCATCTGCTTCGAC 257
 DB 121 CCGGGCTGGGCAACCTGCTGATGCTGACTTCCAGAACACACCATCTGCTTCGAC 180
 QY 258 TTGCGGCGCGCTTCCGGGACGCTTTCAGCTGCTGAGCTGGCTGAGCGCGGTGCTG 317
 DB 181 TTGCGGCGCGCTTCCGGGACGCTTTCAGCTGCTGAGCTGGCTGAGCGCGGTGCTG 240
 QY 318 CTCATATGGCTGGGCGCGCTGGGAGCGCTGGTGAACCCAGCGAGACACCCCGAC 377
 DB 241 CTCATATGGCTGGGCGCGCTGGGAGCGCTGGTGAACCCAGCGAGACACCCCGAC 300
 QY 378 CGCCCGCTGTGCCCATCAACCCAGATCTTGGGCTTTTGGGCGCGCTTCCCAA 428
 DB 301 CGCCCGCTGTGCCCATCAACCCAGATCTTGGGCTTTTGGGCGCGCTTCCCAA 360
 QY 429 ----- 428
 DB 361 CTGGCGCGCTATGGGCCCGCTGGCGGAGCAGAGCGCTTCTCCGCTCTCCACCTTGC 420
 QY 429 ----- 428
 DB 421 AACTTTGGGCTGGGCAAGAAGTCTGCTGGAGCAGTGGGTGACCCGAGGAGCGCGCTT 480

```
QY 429 -----GGAGCGCCCTTTTCGCCCAACGGTCTCTTTGACAAA 464
Db 481 TGTGGCGCTTCGCCCAACCACTCCGGAGCGCCCTTTGCCCAACGGTCTTTTGACAAA 540
QY 465 GCCGTGAGCAACGTGATCGCTCCCTACCTGCGGGCGCGCTTCGAGTACGACACCT 524
Db 541 GCCGTGAGCAACGTGATCGCTCCCTACCTGCGGGCGCGCTTCGAGTACGACACCT 600
QY 525 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGACTGAGAGGAGTCTGGGCTTTCTG 584
Db 601 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGACTGAGAGGAGTCTGGGCTTTCTG 660
QY 585 CGCAGAGTGTCTAAATGCTGTCCCTCTCTGCAATATCCAGAGCGCTGCTGGCAAGTTC 644
Db 661 CGCAGAGTGTCTAAATGCTGTCCCTCTCTGCAATATCCAGAGCGCTGCTGGCAAGTTC 720
QY 645 CTACGCTTCCAAAGGCTTTCTGACCCAGAGTGTGATGCTGCTAACTGAGCACAGATG 704
Db 721 CTACGCTTCCAAAGGCTTTCTGACCCAGAGTGTGATGCTGCTAACTGAGCACAGATG 780
QY 705 ACTCGGACCCAGCCAGCCCGCCCGAGACCTGACTGAGGCTTCTTGGCAGAGATGAG 764
Db 781 ACTCGGACCCAGCCAGCCCGCCCGAGACCTGACTGAGGCTTCTTGGCAGAGATGAG 840
QY 765 AAGGCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACTCGCATAGTGTGGCT 824
Db 841 AAGGCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACTCGCATAGTGTGGCT 900
QY 825 GACCTGTCTCTGCGGGATGTTGACCACTCGACCGCTGGCTGGGGCTCTCTGCTC 884
Db 901 GACCTGTCTCTGCGGGATGTTGACCACTCGACCGCTGGCTGGGGCTCTCTGCTC 960
QY 885 ATGATCTTACATCCGGATGTGAGCGCGTGTCCAAAGAGATCGACGATGTATAGGG 944
Db 961 ATGATCTTACATCCGGATGTGAGCGCGTGTCCAAAGAGATCGACGATGTATAGGG 1020
QY 945 CAGGTGCGCGACACAGAGATGGTGACAGGCTCACATGCCCTCACCACTGCCGTGATT 1004
Db 1021 CAGGTGCGCGACACAGAGATGGTGACAGGCTCACATGCCCTCACCACTGCCGTGATT 1080
QY 1005 CATGAGTGCAGCGCTTTGGGACATCGTCCCTGGGTGTGACCCATATGACATCCCGT 1064
Db 1081 CATGAGTGCAGCGCTTTGGGACATCGTCCCTGGGTGTGACCCATATGACATCCCGT 1140
QY 1065 GACATCGAAGTACAGGCTTCGGATCCCTAAGGGAAAGCACTCATCACCACTGTCA 1124
Db 1141 GACATCGAAGTACAGGCTTCGGATCCCTAAGGGAAAGCACTCATCACCACTGTCA 1200
QY 1125 TCGGTGCTGAAGATGAGCGCTCTGGGAGAGCCCTTCGGCTTCCACCCGAAACACTTC 1184
Db 1201 TCGGTGCTGAAGATGAGCGCTCTGGGAGAGCCCTTCGGCTTCCACCCGAAACACTTC 1260
QY 1185 CTGATGCTCCAGGGCCACTTTGTGAAGCGGAGGCTTCTGCTGCTTCTCAGCAGCGCG 1244
Db 1261 CTGATGCTCCAGGGCCACTTTGTGAAGCGGAGGCTTCTGCTGCTTCTCAGCAGCGCG 1320
QY 1245 CGTGATGCTCGGGAGCCCTTGGCGCGCATGAGGCTTCTCTTCTTACCTCCCTG 1304
Db 1321 CGTGATGCTCGGGAGCCCTTGGCGCGCATGAGGCTTCTCTTCTTCTTCTTCTTCTTCTG 1380
QY 1305 CTGAGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1364
Db 1381 CTGAGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1365 TTTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1424
Db 1441 TTTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
QY 1425 GTACCTAGTCCCGAGCTGCTCCCTAGCAGAGGCTCTTAATGTAACAATAAGCAATGTGG 1484
Db 1501 GTACCTAGTCCCGAGCTGCTCCCTAGCAGAGGCTCTTAATGTAACAATAAGCAATGTGG 1560
```

```
QY 1485 TAGTTCC 1491
Db 1561 TAGTTCC 1567

RESULT 7
ACA61303
ID ACA61303 standard; cDNA; 1567 BP.
XX ACA61303;
AC ACA61303;
XX 16-JUL-2003 (first entry)
XX Human cytochrome p450 gene CYP2D6, wild-type cDNA.
XX Human; ss; gene; cytochrome P450; CYP2D6; chromosome 22; SNP;
XX single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
XX psychiatric disorder; drug sensitivity.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1494
FT /tag= a
FT /product= "CYP2D6 protein"
XX
PN EP1281755-A2.
XX 05-FEB-2003.
XX 16-JUL-2002; 2002EP-00254972.
XX 31-JUL-2001; 2001US-0309111P.
XX (PFIZ ) PFIZER PROD INC.
PI Milos PM, Webb SM;
XX
DR WPI; 2003-373769/36.
DR P-PSDB; AEU09593.
XX
New cytochrome P450 2D6 gene variants and polypeptides, useful for
determining if a subject has or is at risk of developing a drug
sensitivity condition or disorder that is associated with an aberrant
CYP2D6 activity.
Claim 2; Fig 4; 88pp; English.
The invention relates to an isolated nucleic acid comprising a cytochrome
P450 2D6 gene variant, e.g. G5799C or C5816AT (referring to the genomic
sequence or the same variant nucleotide in the corresponding cDNA
sequences). Also included are probes, primers (allele specific
oligonucleotides) and arrays used to detect and or amplify the CYP2D6
gene polymorphic regions, the variant polypeptides, antibodies which are
capable of distinguishing between the variant and wild-type polypeptides,
determining whether a subject has a genetic deficiency for metabolising a
drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
determining whether an individual is susceptible to being a poor
metaboliser of drugs. The DNA probe is useful for hybridising to a
variant form of the CYP2D6 gene. The primer is useful for amplifying the
C5816TA allelic variant. The allele specific nucleotide is useful for the
detection of the C5816TA allelic variant. The methods are useful for
determining whether a subject has a genetic deficiency for metabolising a
drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and
determining if an individual is susceptible to being a poor metaboliser
of drugs. The nucleic acids are useful as probes or primers for
determining whether a subject has a genetic deficiency for metabolising
drugs that are substrates of P450 CYP2D6. The methods are useful for
determining if a subject has or is at risk of developing a drug
sensitivity condition or disorder that is associated with an aberrant
CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
appropriate drugs or determining the course of treatment to administer to
```

CC a subject to treat cardiovascular or psychiatric disorders, or for
CC treating a subject with a drug sensitivity or disorder associated with a
CC specific allelic variant of a polymorphic region of the CYP2D6 gene. The
CC antibodies are useful for monitoring CYP2D6 protein levels in an
CC individual for determining whether a subject has a disease or conditions
CC associated with an aberrant CYP2D6 protein level. The gene is located on
CC human chromosome 22. The present sequence is the wild-type CYP2D6 cDNA
XX
SQ Sequence 1567 BP; 267 A; 527 C; 465 G; 308 T; 0 U; 0 Other;

Query Match 80.6%; Score 1238.2; DB 7; Length 1567;
Best Local Similarity 89.7%; Pred. No. 6e-226; Indels 153; Gaps 1;
Matches 1406; Conservative 0; Mismatches 8;

QY 78 ATGGGGCTAAGACACTGGTGGCCCTGGCGGTGATAGTGCCCATCTTCTGCTCTCTGGTG 137
DB 1 ATGGGGCTAAGACACTGGTGGCCCTGGCGGTGATAGTGCCCATCTTCTGCTCTCTGGTG 60

QY 138 GACCTGATGACACCGGGCCCAACGCTGGGTGACGCTACTCACCAGGCCCTTGGCCACTG 197
DB 61 GACCTGATGACACCGGGCCCAACGCTGGGTGACGCTACTCACCAGGCCCTTGGCCACTG 120

QY 198 CCGGGCTGGGCAACCTGCTGATGAGCTTCCAGAACACACACATCTTCCGACCAAG 257
DB 121 CCGGGCTGGGCAACCTGCTGATGAGCTTCCAGAACACACATCTTCCGACCAAG 180

QY 258 TTGGGGCGCGCTTCCGGGACGCTGTTTCCAGCTGACCTGGCTTGGAGCGCGGTGGTGGT 317
DB 181 TTGGGGCGCGCTTCCGGGACGCTGTTTCCAGCTGACCTGGCTTGGAGCGCGGTGGTGGT 240

QY 318 CTCAATGGGTGGGGCGGTGGCGAGGGCTGGTGAACCCAGCGGAGGACACCGCCGAC 377
DB 241 CTCAATGGGTGGGGCGGTGGCGAGGGCTGGTGAACCCAGCGGAGGACACCGCCGAC 300

QY 378 CGCCCGCTGTGCCATACCCAGATCTTGGGTTTGGCGCGGTGCCAA 428
DB 301 CGCCCGCTGTGCCATACCCAGATCTTGGGTTTGGCGCGGTGCCAA 360

QY 429 ----- 428
DB 361 CTGGCGGCTATGGGCGCGGTGGCGAGGAGGGGCTTCTCCGTGTCACCTTTGGCG 420

QY 429 ----- 428
DB 421 AACTTGGGCTGGGCAAGAGTCCGTGGAGAGTGGGTGACCGAGGAGCGCGCTGCCTT 480

QY 429 -----GGAGCGCCCTTTCGCCCAACCGTCTCTTGGCAAA 464
DB 481 TGTGGCGCTTGCACCAACCACTCCGAGCGCCCTTTCGCCCAACCGTCTCTTGGCAAA 540

QY 465 GCGGTGAGCAACGTGATCGCTCCCTCACCTGGGGCGCGGCTTCCAGTACGACGACCT 524
DB 541 GCGGTGAGCAACGTGATCGCTCCCTCACCTGGGGCGCGGCTTCCAGTACGACGACCT 600

QY 525 CGCTTCTCAGGCTGTGGACCTAGTCTAGGAGGGACTGAAGGAGGAGTGGGCTTTCTG 584
DB 601 CGCTTCTCAGGCTGTGGACCTAGTCTAGGAGGGACTGAAGGAGGAGTGGGCTTTCTG 660

QY 585 CGGAGGTGCTGAATGCTGTCCTCCCTGCTGATATCCAGCGCTGGCTGGCAAGTTC 644
DB 661 CGGAGGTGCTGAATGCTGTCCTCCCTGCTGATATCCAGCGCTGGCTGGCAAGTTC 720

QY 645 CTACGCTTCCAAAGGCTTTCCTGACCCAGCTGGATGAGTCTTAAGTACGACAGATG 704
DB 721 CTACGCTTCCAAAGGCTTTCCTGACCCAGCTGGATGAGTCTTAAGTACGACAGATG 780

QY 705 ACCTGGGACCCAGCCCGCCCGAGACCTGACTGAGGCTTCTTGGCAGAGATGGAG 764
DB 781 ACCTGGGACCCAGCCCGCCCGAGACCTGACTGAGGCTTCTTGGCAGAGATGGAG 840

QY 765 AAGGCCAAGGGGAACCTTGAGGACGCTTCAATGATGAGAACCTGGCGATAGTGGTGGCT 824
DB 841 AAGGCCAAGGGGAACCTTGAGGACGCTTCAATGATGAGAACCTGGCGATAGTGGTGGCT 900

QY 825 GACCTGTTCTCTGCGGGATGTGACCACTGACACGCTGGCTGGGCTCTCTGCTC 884
DB 901 GACCTGTTCTCTGCGGGATGTGACCACTGACACGCTGGCTGGGCTCTCTGCTC 960

QY 885 ATGATCTACATCCGGATGTGACGCGCGTGTCCAAACAGGAGATCGACGACGTGATGG 944
DB 961 ATGATCTACATCCGGATGTGACGCGCGTGTCCAAACAGGAGATCGACGACGTGATGG 1020

QY 945 CAGGTGCGGCGACACAGAGATGGGTGACAGGCTCAGCTGCTCACCACCTGCGGTGATT 1004
DB 1021 CAGGTGCGGCGACACAGAGATGGGTGACAGGCTCAGCTGCTCACCACCTGCGGTGATT 1080

QY 1005 CATGAGGTGACGCGCTTTGGGGACATCGTCCCTCGGTGGTGAACCATATGACATCCCT 1064
DB 1081 CATGAGGTGACGCGCTTTGGGGACATCGTCCCTCGGTGGTGAACCATATGACATCCCT 1140

QY 1065 GACATCGAGTACAGGGCTTCCGATCCCTTAGGGAACGACACTCATCACCACCTGTCA 1124
DB 1141 GACATCGAGTACAGGGCTTCCGATCCCTTAGGGAACGACACTCATCACCACCTGTCA 1200

QY 1125 TCGGTGCTGAAGGATGAGGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCGAACTTC 1184
DB 1201 TCGGTGCTGAAGGATGAGGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCGAACTTC 1260

QY 1185 CTGGATGCCAGGGCCACTTTGTGAAGCGGAGGCTTCTCTGCTTTCTCAGCAGGCGC 1244
DB 1261 CTGGATGCCAGGGCCACTTTGTGAAGCGGAGGCTTCTCTGCTTTCTCAGCAGGCGC 1320

QY 1245 CGTGATGCTCGGGAGGCGCTTGGCGGATGAGAGCTTCTCTTCTTACCTCCCTG 1304
DB 1321 CGTGATGCTCGGGAGGCGCTTGGCGGATGAGAGCTTCTCTTCTTACCTCCCTG 1380

QY 1305 CTGACGACTTCCAGCTTCTCGGTGCTTCTGAGAGCCCGGCGGAGGCTTCTGCTGTC 1364
DB 1381 CTGACGACTTCCAGCTTCTCGGTGCTTCTGAGAGCCCGGCGGAGGCTTCTGCTGTC 1440

QY 1365 TTTGCTTCTGCTGAGCCCACTCCCTTATGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1424
DB 1441 TTTGCTTCTGCTGAGCCCACTCCCTTATGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1500

QY 1425 GTACCTAGTCCCGAGGCTGCTCCCTAGCCAGAGGCTCTAATGATACAAATAAGCAATGTCG 1484
DB 1501 GTACCTAGTCCCGAGGCTGCTCCCTAGCCAGAGGCTCTAATGATACAAATAAGCAATGTCG 1560

QY 1485 TAGTTCC 1491
DB 1561 TAGTTCC 1567

RESULT 8

AAQ12893
ID AAQ12893 standard; cDNA; 1586 BP.
XX
AC AAQ12893;
XX AC
XX 25-MAR-2003 (revised)
DT 22-OCT-1991 (first entry)
XX
DE Debrisoquine hydroxylase variant a.
XX db1; Extensive metaboliser/poor metaboliser; EM/PM; ss.
XX Synthetic.
XX
FH Key
FT exon
FT 1. .180
FT /*tag= a
FT /number= 1
FT intron
FT 181. .245
FT /*tag= b
FT /number= 1
FT /note= "first 64 bases of intron 1"


```
FT conflict 383
FT /*tag= c
FT /note= "c.f. C(383) in Gonzalez et al., (1988) Nature,
FT 331, 442-445."
FT mutation 506
FT /*tag= d
FT /note= "G(506) is deleted; results in frameshift c.f.
FT genomic P450I12D sequence"
XX WO9110745-A.
XX
XX 25-JUL-1991.
XX
XX 18-JAN-1990; 90GB-00001181.
XX
XX 18-JAN-1990; 90GB-00001181.
XX (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
XX
XX Wolf CR, Miles JS, Spurr NK, Gough AC;
XX WPI; 1991-238033/32.
XX
XX Identification of cytochrome P450 enzyme bufuralol-1'-hydroxylase - as
XX indication of extensive-poor metaboliser pheno:types, important for drug
XX dosage.
XX
XX Example 1; Fig 2; 43pp; English.
XX
XX This full-length dbl variant sequence was obtained from two overlapping
XX clones (pMP32 and pMP33). It is not expected to encode a functional
XX protein as it contains part of intron 1 and has a frameshift relative to
XX the functional gene. The other base pair differences between the
XX pMP32/pMP33 sequence and the normal dbl sequence form the basis of
XX detection methods designed to distinguish EM/PM phenotypes. For example,
XX at position 294 of the variant sequence G replaces the wild-type C. This
XX introduces a HaeIII restriction site in PM individuals, PCR amplification
XX of a 172bp fragment containing this site, followed by HaeIII digestion
XX allows homozygous EM individuals (who lack this site) to be distinguished
XX from PM individuals. See AAQ12885-Q12894. (Updated on 25-MAR-2003 to
XX correct PA field.)
XX
XX Sequence 1586 BP; 286 A; 527 C; 464 G; 309 T; 0 U; 0 Other;
XX
XX Query Match 80.3%; Score 1234.6; DB 2; Length 1586;
XX Best Local Similarity 89.7%; Pred. No. 2.9e-225;
XX Matches 1424; Conservative 0; Mismatches 9; Indels 154; Gaps 3;
XX
XX 78 ATGGGGCTAGAGACACTGGTCCCTGGCCGTGATAGTGGCCATCTTCTGCTCTGCTG 137
XX 1 ATGGGGCTAGAGACACTGGTCCCTGGCCGTGATAGTGGCCATCTTCTGCTCTGCTG 60
XX
XX 138 GACCTGATGACCGCGCCCAAGCTGGGCTGCACGCTACTACACAGGCCCTCTGCCACTG 197
XX
XX 61 GACCTGATGACCGCGCCCAAGCTGGGCTGCACGCTACTACACAGGCCCTCTGCCACTG 120
XX
XX 198 CCCGGCTGGGCAACTGCTGATGCTGACCTTCAGAAACACACATCTGCTTGACACAG 257
XX
XX 121 CCCGGCTGGGCAACTGCTGATGCTGACCTTCAGAAACACACATCTGCTTGACACAG 180
XX
XX 258 TTGGGGCGCGCTTCGGGAGCTGTTACGCTGACGCTGGCTGGACCGCCGGTGGTGGT 317
XX
XX 181 TGTTCCGGCCGCTTCGGGAGCTGTTACGCTGACGCTGGCTGGACCGCCGGTGGTGGT 240
XX
XX 318 CTCAATGGGCTGGCGCCCGTGGCGAGCGCTGGTGGACCCACGGCGAGGACACCGCCGAC 377
XX
XX 241 CTCAATGGGCTGGCGCCCGTGGCGAGCGCTGGTGGACCCCGGGCGAGGACACCGCCGAC 300
XX
XX 378 CGCCCGCTGTGCCCATACCCAGATCTCTGGTGGTGGTGGCGCGCTTCCCAAG----- 429
XX
XX 301 CGCCCGCTGTGCCCATACCCAGATCTCTGGTGGTGGTGGCGCGCTTCCCAAGGGGTGTC 360
XX
XX 430 ----- 429
```

```
Db 361 CTGGGGCTATGGGCGCGCTGGCGGACGAGAGGCGCTTCTCGGTCTCCACCTTGCGC 420
Qy 430 ----- 429
Db 421 AACTGGGCTGGGCAAGAGTCTGAGAGTGGGTGACCGAGGAGCGCGCTGCGCTT 480
Qy 430 -----GAGCCCTTTTGGCCCCCAAGGCTCTTTGGACAAAG 465
Db 481 TGTGCCGCTTTCGCAACCACTCCGACGCGCTTTTGGCCCCCAAGGCTCTTTGGACAAAG 540
Qy 466 CCGTGAGCAACCTGATCGCCCTCCCTCACCTTCGGGCGCGCTTTCGAGTACGACCACTC 525
Db 541 CCGTGAGCAACCTGATCGCCCTCCCTCACCTTCGGGCGCGCTTTCGAGTACGACCACTC 600
Qy 526 GCTTCCTCAGGCTGCTGGAACCTAGCTCAGGAGGACTGAAGAGGAGTTCGGGCTTTTCG 585
Db 601 GCTTCCTCAGGCTGCTGGAACCTAGCTCAGGAGGACTGAAGAGGAGTTCGGGCTTTTCG 660
Qy 586 GCGAGGTGCTGAATGCTGCTCCCGTCTCCCTGCATATCCAGCGCTGGCTGGCAAGTCC 645
Db 661 GCGAGGTGCTGAATGCTGCTCCCGTCTCCCTGCATATCCAGCGCTGGCTGGCAAGTCC 720
Qy 646 TACGCTTCCAAAGGCTTTCTGTACCCAGCTGGATGAGCTGCTAACTAGCAGCAGGATGA 705
Db 721 TACGCTTCCAAAGGCTTTCTGTACCCAGCTGGATGAGCTGCTAACTAGCAGCAGGATGA 780
Qy 706 CTGGGACCCAGCCAGCCCGCCGAGACCTGACTGAGGCGCTTCTGGCAGAGATGGAGA 765
Db 781 CTGGGACCCAGCCAGCCCGCCGAGACCTGACTGAGGCGCTTCTGGCAGAGATGGAGA 840
Qy 766 AGGCCAAGGGGAAACCTGAGAGGAGCTTCAATGATGAGAACCTGGCATAGTGGCTG 825
Db 841 AGGCCAAGGGGAAACCTGAGAGGAGCTTCAATGATGAGAACCTGGCATAGTGGCTG 900
Qy 826 ACCTGTTCTCTGCGGAGTGGTACCACTCGACACGCTGGCGCTGGCGCTCTCTGCTCA 885
Db 901 ACCTGTTCTCTGCGGAGTGGTACCACTCGACACGCTGGCGCTGGCGCTCTCTGCTCA 960
Qy 886 TGATCTTACATTCGAGATGTGACGCGCTGTCACAGGAGATCGACAGCT -GATAGGG 944
Db 961 TGATCTTACATTCGAGATGTGACG -CGGTGTCCAAACAGGAGATCGACAGCTGATAGG 1019
Qy 945 CAGGTGCGGACACAGAGATGGGTGACAGGCTCACATGCCCTACACACTGCGCTGATT 1004
Db 1020 CAGGTGCGGACACAGAGATGGGTGACAGGCTCACATGCCCTACACACTGCGCTGATT 1079
Qy 1005 CATGAGTGCAGCGCTTTGGGGACATCGTCCCCCTGGGTGTGACCCATATGACATCCCGT 1064
Db 1080 CATGAGTGCAGCGCTTTGGGGACATCGTCCCCCTGGGTGTGACCCATATGACATCCCGT 1139
Qy 1065 GACATCAAGTGCAGCGCTTTGGGGACATCGTCCCCCTGGGTGTGACCCATATGACATCCCGT 1124
Db 1140 GACATCAAGTGCAGCGCTTTGGGGACATCGTCCCCCTGGGTGTGACCCATATGACATCCCGT 1199
Qy 1125 TCGGTGCTGAGGATGAGGCGCTCTGGAGAGAGCCCTTCGCTTCCACCCGACACTTC 1184
Db 1200 TCGGTGCTGAGGATGAGGCGCTCTGGAGAGAGCCCTTCGCTTCCACCCGACACTTC 1259
Qy 1185 CTGATGCCACAGGCGCACTTTGTGAAGCGGAGGCGCTTCTGCTTCTCAGCAGGCGCG 1244
Db 1260 CTGATGCCACAGGCGCACTTTGTGAAGCGGAGGCGCTTCTGCTTCTCAGCAGGCGCG 1319
Qy 1245 CGTGCATGCTTCGGGAGCGCTTCGGCCCGATGAGGCTTCTCTCTTCTTCACTCCCTG 1304
Db 1320 CGTGCATGCTTCGGGAGCGCTTCGGCCCGATGAGGCTTCTCTCTTCTTCACTCCCTG 1379
Qy 1305 CTGAGCACTTCAGTTCCTGGTGGCCCACTGGAGAGCGCGCGCCGACGACCACTGCTGTC 1364
Db 1380 CTGAGCACTTCAGTTCCTGGTGGCCCACTGGAGAGCGCGCGCCGACGACCACTGCTGTC 1439
Qy 1365 TTTGCTTCTCTGCTGAGCCCATCCCTCATGAGCTTGTGTGCTGTCGCCCGCTAGATGGG 1424
```

Db 1440 TTGCTTTCTGGTGACCCATCCCTATGAGCTTTGCTGTGCCCCCTAGAAATGGG 1499
Qy 1425 GTACCTAGTCCCGACGCTGCTCCCTAGCCAGAGGCTCTAATGTACAAATGAAGCAATGTGG 1484
Db 1500 GTACCTAGTCCCGACGCTGCTCCCTAGCCAGAGGCTCTAATGTACAAATGAAGCAATGTGG 1559
Qy 1485 TAGTTCACAAAAAATTT 1511
Db 1560 TAGTTCACAAAAAATTT 1586

RESULT 9
ACA61304
ID ACA61304 standard; cDNA; 1568 BP.
XX AC ACA61304;
XX AC ACA61304;
XX 16-JUL-2003 (first entry)
XX Human cytochrome p450 gene CYP2D6 C5816TA variant, cDNA.
XX Human; ss; gene; cytochrome P450; CYP2D6; chromosome 22; SNP;
KW single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
KW psychiatric disorder; drug sensitivity.
XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 1..1495
XX FT /*tag= a
XX FT /product= "CYP2D6 protein"
XX FT replace(1474..1475,C)
XX FT /*tag= b
XX FT /standard_name= "Single nucleotide polymorphism"
XX EP1281755-A2.
XX 05-FEB-2003.
XX 16-JUL-2002; 2002EP-00254972.
XX 31-JUL-2001; 2001US-030911P.
XX (PFIZ) PFIZER PROD INC.
XX Milos PM, Webb SM;
XX WPI; 2003-373769/36.
XX P-PSDB; ABU09594.
XX
XX New cytochrome P450 2D6 gene variants and polypeptides, useful for
PT determining if a subject has or is at risk of developing a drug
PT sensitivity condition or disorder that is associated with an aberrant
PT CYP2D6 activity.
XX
XX Claim 3; Fig 5; 88pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising a cytochrome
CC P450 2D6 gene variant, e.g. G5799C or C5816AT (referring to the genomic
CC sequence or the same variant nucleotide in the corresponding cDNA
CC sequences). Also included are probes, primers (allele specific
CC oligonucleotides) and arrays used to detect and/or amplify the CYP2D6
CC gene polymorphic regions, the variant polypeptides, antibodies which are
CC capable of distinguishing between the variant and wild-type polypeptides,
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
CC determining whether an individual is susceptible to being a poor
CC metaboliser of drugs. The DNA probe is useful for hybridising to a
CC variant form of the CYP2D6 gene. The primer is useful for amplifying the
CC C5816TA allelic variant. The allele specific nucleotide is useful for the
CC detection of the C5816TA allelic variant. The methods are useful for
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and

CC determining if an individual is susceptible to being a poor metaboliser
CC of drugs. The nucleic acids are useful as probes or primers for
CC determining whether a subject has a genetic deficiency for metabolising
CC drugs that are substrates of P450 CYP2D6. The methods are useful for
CC determining if a subject has or is at risk of developing a drug
CC sensitivity condition or disorder that is associated with an aberrant
CC CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
CC aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
CC appropriate drugs or determining the course of treatment to administer to
CC a subject to treat cardiovascular or psychiatric disorders, or for
CC treating a subject with a drug sensitivity or disorder associated with a
CC specific allelic variant of a polymorphic region of the CYP2D6 gene. The
CC antibodies are useful for monitoring CYP2D6 protein levels in an
CC individual for determining whether a subject has a disease or condition
CC associated with an aberrant CYP2D6 protein level. The gene is located on
CC human chromosome 22. The present sequence is the CYP2D6 cDNA carrying the
CC C5816TA variation
XX
SQ Sequence 1568 BP; 268 A; 526 C; 465 G; 309 T; 0 U; 0 Other;
Query Match 79.7%; Score 1225.6; DB 7; Length 1568;
Best Local Similarity 89.6%; Pred. No. 1.5e-223;
Matches 1405; Conservative 0; Mismatches 9; Indels 154; Gaps 2;
Qy 78 ATGGGGCTAGAAGCACTGCTGCCCTGGCGCTGATGTGGCCATCTTCCTGCTCTGGTG 137
Db 1 ATGGGGCTAGAAGCACTGCTGCCCTGGCGCTGATGTGGCCATCTTCCTGCTCTGGTG 60
Qy 138 GACCTGATGACGGCGGCCAACGCTGGGCTGCACGCTACTCACCAGGCCCTCCCTGCACATG 197
Db 61 GACCTGATGACGGCGGCCAACGCTGGGCTGCACGCTACTCACCAGGCCCTCCCTGCACATG 120
Qy 198 CCGGGCTGGGCAACCTGCTGCATGTGGACTTCCAGAACACACCATCTGCTTCGACCCAG 257
Db 121 CCGGGCTGGGCAACCTGCTGCATGTGGACTTCCAGAACACACCATCTGCTTCGACCCAG 180
Qy 258 TTGCGGCGCCCTTCGGGCACTGTTGACCTGCACTGGCTGGACCGCGGTGGTCTGCTG 317
Db 181 TTGCGGCGCCCTTCGGGCACTGTTGACCTGCACTGGCTGGACCGCGGTGGTCTGCTG 240
Qy 318 CTCATGTTGGCTGGCGCGCTGGCGAGGGCTGGTGGACCCAGCGGAGAGACACCGCGGAC 377
Db 241 CTCATGTTGGCTGGCGCGCTGGCGAGGGCTGGTGGACCCAGCGGAGAGACACCGCGGAC 300
Qy 378 CGCCCGCTGTGCCCATCACCATCCTGCGGTGTTGGCGCGCTGTTCCCAA----- 428
Db 301 CGCCCGCTGTGCCCATCACCATCCTGCGGTGTTGGCGCGCTGTTCCCAAAGGGGTGTT 360
Qy 429 ----- 428
Db 361 CTGGCGCGCTATGGGCGCGCTGGCGAGCAGAGCGGCTTCTCCGTGTCCACCTTTCGCG 420
Qy 429 ----- 428
Db 421 AACTTGGGCTGGGCAAGAGTCTGTTGGAGCAGTGGTGACCGAGAGGAGCGCTTCCTT 480
Qy 429 -----GGAGCGCCCTTTTGGCCCAACCGTCTCTTTGGACAAA 464
Db 481 TGTGGCGCTTCGCCCAACCACTCCGAGCGCCCTTTTGGCCCAACCGTCTCTTTGGACAAA 540
Qy 465 GCCGTGAGCAACGTGATCGCTCCCTCCTCAGCTGGGCGCGCTTCGAGTACGACGACCTT 524
Db 541 GCCGTGAGCAACGTGATCGCTCCCTCCTCAGCTGGGCGCGCTTCGAGTACGACGACCTT 600
Qy 525 CGCTTCTCAGGCTGCTGGAACCTAGCTCAGAGGAGCTGAGAGGAGGAGTGGGGCTTTCG 584
Db 601 CGCTTCTCAGGCTGCTGGAACCTAGCTCAGAGGAGGAGTGAAGAGGAGGAGTGGGGCTTTCG 660
Qy 585 CGCAGGAGTCTGAATGCTGCTCCCTCCTCTGATATCCAGCGCTGCTGCGAGGTC 644
Db 661 CGCAGGAGTCTGAATGCTGCTCCCTCCTCTGATATCCAGCGCTGCTGCGAGGTC 720
Qy 645 CTACGCTTCCAAAAGGCTTTCTCTGACCCAGCTGGATGAGCTGCTAACTGAGCACGAGTATG 704

Db 721 CTACGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGTCTAACTGAGCACAGGATG 780
QY 705 ACCTGGGACCCAGCCAGCCCGCCGAGACCTGACTGAGGCTTCTCTGGCAGAGATGAG 764
Db 781 ACCTGGGACCCAGCCAGCCCGCCGAGACCTGACTGAGGCTTCTCTGGCAGAGATGAG 840
QY 765 AAGGCCAAGGGGAAACCTCTGAGAGAGCTTCAATGATGAGAACCTGCGCATAGTGGTGGCT 824
Db 841 AAGGCCAAGGGGAAACCTCTGAGAGAGCTTCAATGATGAGAACCTGCGCATAGTGGTGGCT 900
QY 825 GACCTGTTCTCTGCGGGATGTCACCACTCGACCGCTGCGCTGGGGCTCTCTGCTC 884
Db 901 GACCTGTTCTCTGCGGGATGTCACCACTCGACCGCTGCGCTGGGGCTCTCTGCTC 960
QY 885 ATGATCTTACATCCGATGTCAGCGCGCTGTCCAAAGAGAGATCGACGATGATAGG 944
Db 961 ATGATCTTACATCCGATGTCAGCGCGCTGTCCAAAGAGAGATCGACGATGATAGG 1020
QY 945 CAGGTGCGGACACAGATGGTGACAGGCTACATGCGCTACCACTGCGCTGAT 1004
Db 1021 CAGGTGCGGACACAGATGGTGACAGGCTACATGCGCTACCACTGCGCTGAT 1080
QY 1005 CATGAGGTGACGCGCTTTGGGACATCGTCCCGCTGGGTGTGACCCATATGACATCCCGT 1064
Db 1081 CATGAGGTGACGCGCTTTGGGACATCGTCCCGCTGGGTGTGACCCATATGACATCCCGT 1140
QY 1065 GACATCGAAGTACAGGCTTCCGATTCCTTAAGGAACGACACTCATCAACAACTGTCA 1124
Db 1141 GACATCGAAGTACAGGCTTCCGATTCCTTAAGGAACGACACTCATCAACAACTGTCA 1200
QY 1125 TCGGTGCTGAGGATGAGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCGACACTTC 1184
Db 1201 TCGGTGCTGAGGATGAGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCGACACTTC 1260
QY 1185 CTGATGCGCCAGGGCCATTTTGTGAAGCGGAGGCTTCTCGCTTTCTCAGCAGGCGC 1244
Db 1261 CTGATGCGCCAGGGCCATTTTGTGAAGCGGAGGCTTCTCGCTTTCTCAGCAGGCGC 1320
QY 1245 CQTGATGCTCGGGAGCCCTGCGCGCATGAGAGCTTCTCTTCTTCTTCTTCTTCTTCTG 1304
Db 1321 CQTGATGCTCGGGAGCCCTGCGCGCATGAGAGCTTCTCTTCTTCTTCTTCTTCTTCTG 1380
QY 1305 CTGACGACTTCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1364
Db 1381 CTGACGACTTCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1365 TTTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1423
Db 1441 TTTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
QY 1424 GGTACTAGTCCCGAGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 1483
Db 1501 GGTACTAGTCCCGAGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 1560
QY 1484 GTAGTTC 1491
Db 1561 GTAGTTC 1568

RESULT 10

ID ACA61305 standard; cDNA; 1565 BP.
XX ACA61305;
XX ACA61305;
DT 16-JUL-2003 (first entry)
XX Human cytochrome p450 gene CYP2D6 G5799C/C58167A variant, cDNA.
XX Human; ss; gene; cytochrome P450; CYP2D6; chromosome 22; SNP;
KW single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
KW psychiatric disorder; drug sensitivity.

XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 1..1506
FT /*tag= a
FT /product= "CYP2D6 protein"
FT /partial
FT /note= "No stop codon shown"
FT replace(1457,G)
FT /*tag= b
FT variation
FT
FT variation
FT replace(1474..1475,C)
FT /*tag= C
FT /standard_name= "Single nucleotide polymorphism"
XX EP1281755-A2.
XX
XX 05-FEB-2003.
XX
XX 16-JUL-2002; 2002EP-00254972.
XX
XX 31-JUL-2001; 2001US-0309111P.
XX (PFIZ) PFIZER PROD INC.
XX
XX Milos PM, Webb SM;
XX
XX WPI; 2003-373769/36.
XX P-PSDB; ABU09595.
XX
XX New cytochrome P450 2D6 gene variants and polypeptides, useful for
XX determining if a subject has or is at risk of developing a drug
XX sensitivity condition or disorder that is associated with an aberrant
XX CYP2D6 activity.
XX
XX Disclosure; Fig 8; 88pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising a cytochrome
XX P450 2D6 gene variant, e.g. G5799C or C58167A (referring to the genomic
XX sequence) or the same variant nucleotide in the corresponding cDNA
XX sequences). Also included are probes, primers (allele specific
XX oligonucleotides) and arrays used to detect and or amplify the CYP2D6
XX gene polymorphic regions, the variant polypeptides, antibodies which are
XX capable of distinguishing between the variant and wild-type polypeptides,
XX determining whether a subject has a genetic deficiency for metabolising a
XX drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
XX determining whether an individual is susceptible to being a poor
XX metaboliser of drugs. The DNA probe is useful for hybridising to a
XX variant form of the CYP2D6 gene. The primer is useful for amplifying the
XX C58167A allelic variant. The allele specific nucleotide is useful for the
XX detection of the C58167A allelic variant. The methods are useful for
XX determining whether a subject has a genetic deficiency for metabolising a
XX drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and
XX determining if an individual is susceptible to being a poor metaboliser
XX of drugs. The nucleic acids are useful as probes or primers for
XX determining whether a subject has a genetic deficiency for metabolising
XX drugs that are substrates of P450 CYP2D6. The methods are useful for
XX determining if a subject has or is at risk of developing a drug
XX sensitivity condition or disorder that is associated with an aberrant
XX CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
XX aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
XX appropriate drugs or determining the course of treatment to administer to
XX a subject to treat cardiovascular or psychiatric disorders, or for
XX treating a subject with a drug sensitivity or disorder associated with a
XX specific allelic variant of a polymorphic region of the CYP2D6 gene. The
XX antibodies are useful for monitoring CYP2D6 protein levels in an
XX individual for determining whether a subject has a disease or conditions
XX associated with an aberrant CYP2D6 protein level. The gene is located on
XX human chromosome 22. The present sequence is the CYP2D6 cDNA carrying the
XX G5799C/C58167A double variation
XX
XX Sequence 1565 BP; 267 A; 527 C; 463 G; 308 T; 0 U; 0 Other;

Query Match			
Best Local Similarity 89.5%; Pred. No. 8.2e-221;			
Matches 1403; Conservative 0; Mismatches 8; Indels 157; Gaps 3;			
QY	78	ATGGGGCTAGAACACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCTGCTCTCTGGTG	137
DB	1	ATGGGGCTAGAACACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCTGCTCTCTGGTG	60
QY	138	GACCTGATGACACCGGCGCAACGCTGGGTGACGCTACTCACCAGGCCCCCTGCCACTG	197
DB	61	GACCTGATGACACCGGCGCAACGCTGGGTGACGCTACTCACCAGGCCCCCTGCCACTG	120
QY	198	CCCGGGCTGGCAACCTGCTGATGGAATTTCCAGAACACACCACTACTGTTCCAGCCAG	257
DB	121	CCCGGGCTGGCAACCTGCTGATGGAATTTCCAGAACACACCACTACTGTTCCAGCCAG	180
QY	258	TTGCGGCGCGCTTCCGGGACGTTTTCAGCTGACGCTGGCTGGACGCGCGGTGCTG	317
DB	181	TTGCGGCGCGCTTCCGGGACGTTTTCAGCTGACGCTGGCTGGACGCGCGGTGCTG	240
QY	318	CTCAATGGGCTGGCGCGCTGGCGAGGCGTGGTGACCCACGCGGAGGACACCGCCGAC	377
DB	241	CTCAATGGGCTGGCGCGCTGGCGAGGCGTGGTGACCCACGCGGAGGACACCGCCGAC	300
QY	378	CGCCCGCTGTGCCCATCACCAGATCTCTGGGTTTGGGCGCGCTTCCCAA	428
DB	301	CGCCCGCTGTGCCCATCACCAGATCTCTGGGTTTGGGCGCGCTTCCCAA	360
QY	429	-----	428
DB	361	CTGGCGCGCTATGGGCGCGGTGGCGGCGAGGCGCTTCTCCGTGCCACCTTTGGCG	420
QY	429	-----	428
DB	421	AACCTGGGCTGGGCAAGAAAGTGGTGGAGCGAGTGGGTGACCGAGAGGCGCGCTT	480
QY	429	-----GGACGCGCCCTTTTCGCCCCCAACCGGTCTCTTTGACAA	464
DB	481	TGTGGCGCTTCCGCAACCACTCCGAGCGCCCTTTCCGCCCAACCGGTCTCTTTGACAA	540
QY	465	GCGTGAGCAACCTGATGCTCCCTCAGCTGGGCGCGCTTCAGTACGACGACCT	524
DB	541	GCGTGAGCAACCTGATGCTCCCTCAGCTGGGCGCGCTTCAGTACGACGACCT	600
QY	525	CGCTTCTCAGGCTGCTGACCTGCTCAGGAGGAGTGAAGGAGGAGTGGGCTTCTG	584
DB	601	CGCTTCTCAGGCTGCTGACCTGCTCAGGAGGAGTGAAGGAGGAGTGGGCTTCTG	660
QY	585	CGGAGGTGTAATGCTGTCCCGCTCTCTGCATATCCAGCGCTGGCTGGCAAGTTC	644
DB	661	CGGAGGTGTAATGCTGTCCCGCTCTCTGCATATCCAGCGCTGGCTGGCAAGTTC	720
QY	645	CTACGCTTCAAAAGGCTTTCCTGAACCGAGTGGATGAGTGTAACTGAGCACAGGATG	704
DB	721	CTACGCTTCAAAAGGCTTTCCTGAACCGAGTGGATGAGTGTAACTGAGCACAGGATG	780
QY	705	ACTTGGGACCCAGCCCGCCCGAGACCTGACTGAGGCTTCTGGCAGAGTGGAG	764
DB	781	ACTTGGGACCCAGCCCGCCCGAGACCTGACTGAGGCTTCTGGCAGAGTGGAG	840
QY	765	AAGGCCAAGGGGAACTCTGAGACGCTTCAATGATGAGAACTCTGGCATAGTGGTGGCT	824
DB	841	AAGGCCAAGGGGAACTCTGAGACGCTTCAATGATGAGAACTCTGGCATAGTGGTGGCT	900
QY	825	GACCTGTTCTTCCCGGGATGGTGACCACTCGACCACTGGCTGGCGCTTCTGCTC	884
DB	901	GACCTGTTCTTCCCGGGATGGTGACCACTCGACCACTGGCTGGCGCTTCTGCTC	960
QY	885	ATGATCTTACATCCGATGTGACGCGCGTGTCCAAACAGGAGATCAACGCTGATAGG	944
DB	961	ATGATCTTACATCCGATGTGACGCGCGTGTCCAAACAGGAGATCAACGCTGATAGG	1020

RESULT 11

AAV19496

ID AAV19496 standard; DNA; 1545 BP.

AC AAV19496;

XX AAV19496;

XX 18-SEP-1998 (first entry)

DE Cytochrome P450IId6 encoding DNA.

DE Human; cytochrome P450IId6; liver/kidney microsomes; LKM-1; antibody;

KW autoimmunohepatitis type II; AIH type II; hepatitis C virus; ss.

OS Hepatitis C virus.

XX Key

XX Location/Qualifiers

FT 1..1494

FT /tag= a

FT /product= "cytochrome P450IId6"

FT /transl_except= (pos:193..195,aa:Phe)

FT /transl_except= (pos:196..198,aa:Gly)

FT /transl_except= (pos:214..216,aa:Gln)

FT /transl_except= (pos:859..861,aa:Glu)

FT /transl_except= (pos:1120..1122,aa:Met)

FT /transl_except= (pos:1162..1164,aa:Arg)

XX JP10101698-A.

XX 21-APR-1998.

XX 24-SEP-1996; 96JP-00273055.

XX

PR 24-SEP-1996; 96JP-00273055.
XX (HOKE-) HOKEN KAGAKU KENKYUSHO KK.
XX
DR WPI; 1998-292131/26.
XX P-PSDB; AAW44869.
XX
PT Amino acid sequence for liver/kidney microsome (LKM) 1 antibody - to
PT human auto-immuno-hepatitis (AIH) type II and hepatitis C.
XX
PS Disclosure; Fig 5; 9pp; Japanese.
XX
CC The present sequence encodes the protein of cytochrome P450IId6.
CC Specifically claimed regions of cytochrome P450IId6 are recognised by AIH
CC Type IIA/IIb LKM1 antibody, see AAW44870 and AAW44871 (encoded by
CC AAV19497 and AAV19498). The antigens and their complementing antibodies
CC are used in the detection of auto-immuno-hepatitis type II and hepatitis
CC C
XX
SQ Sequence 1545 BP; 257 A; 525 C; 457 G; 301 T; 0 U; 5 Other;
Query Match 78.1%; Score 1201; DB 2; Length 1545;
Best Local Similarity 88.9%; Pred. No. 7.1e-219;
Matches 1373; Conservative 0; Mismatches 19; Indels 153; Gaps 1;
QY 78 ATGGGGCTAGAGCACTGGTCCCTGGCGCTGATAGTGGCCATCTTCTGCTCTGCTG 137
DB 1 ATGGGGCTAGAGCACTGGTCCCTGGCGCTGATAGTGGCCATCTTCTGCTCTGCTG 60
QY 138 GACCTGATGACACCGCGCGCAAGCTGGGCTGACGCTACTACCAAGGCGCCCTGCACTG 197
DB 61 GACCTGATGACACCGCGCGCAAGCTGGGCTGACGCTACTACCAAGGCGCCCTGCACTG 120
QY 198 CCCGGCTGGGCAACCTGCTGATGTGACTTTCAGAAACACACATATGCTTGCACAG 257
DB 121 CCCGGCTGGGCAACCTGCTGATGTGACTTTCAGAAACACACATATGCTTGCACAG 180
QY 258 TTGCGCGCGCGCTTCGGGAGCTGTTACGCTGACGCTGAGCTGGCTGAGCGCGGTGCTG 317
DB 181 TTGCGCGCGCGCTTCGGGAGCTGTTACGCTGACGCTGAGCTGGCTGAGCGCGGTGCTG 240
QY 318 CTCATATGGGCTGGCGCGCGCTGGCGAGGCGTGTGACCCAGCGGAGGACACCGCGAC 377
DB 241 CTCATATGGGCTGGCGCGCGCTGGCGAGGCGTGTGACCCAGCGGAGGACACCGCGAC 300
QY 378 CGCCCGCTGTGCCCATACCCAGATCTTGGGTTTGGGCGCGGTTCCTCAA----- 428
DB 301 CGCCCGCTGTGCCCATACCCAGATCTTGGGCTTCGGGCGCGGTTCCTCAAAGGGGTGTC 360
QY 429 ----- 428
DB 361 CTGGCGCGCTATGGCGCGCGTGGCGGAGCAGAGGCGCTTCTCGTCTCCACCTTGGCG 420
QY 429 ----- 428
DB 421 AACTTGGGCTGGGCAAGAAGTCTGCTGGAGCAGTGGGTGACCGAGGCGNGCGCTT 480
QY 429 -----GGAGCGCCCTTTTGGCGCGCGCGCTTCTTGGACAAA 464
DB 481 TGTGCGCGCTTGGCAACCACTCCCGAGCGCGCTTTTGGCGCGCGCTTCTTGGACAAA 540
QY 465 GCGGTGAGCAAGTATGCGCTCCCTACCTGCGGCGCGCTTTCGAGTACGACGCCCT 524
DB 541 GCGGTGAGCAAGTATGCGCTCCCTACCTGCGGCGCGCTTTCGAGTACGACGCCCT 600
QY 525 CGCTTCTCAGGCTCTGACCTAGCTCAGGAGGAGTGAAGGAGGAGTCTGGGCTTCTG 584
DB 601 CGCTTCTCAGGCTCTGACCTAGCTCAGGAGGAGTGAAGGAGGAGTCTGGGCTTCTG 660
QY 585 CCGGAGGTGCTGAATGCTGCTCCCTCTCTGATATCCAGCGCTGGTGGGAGGTC 644
DB 661 CCGGAGGTGCTGAATGCTGCTCCCTCTCTGATATCCCGCGCTGGTGGGAGGTC 720

QY 645 CTAGCTTCCAAAAGGCTTTCTCACCAGCTGGATGAGCTGCTAACTAGCAGCAGGATG 704
DB 721 CTAGCTTCCAAAAGGCTTTCTCACCAGCTGGATGAGCTGCTAACTAGCAGCAGGATG 780
QY 705 ACCTGGACACCGACCCAGCCCCCGAGACCTGACTGAGGCTTCTTGGGAGAGATGAG 764
DB 781 ACCTGGACACCGACCCAGCCCCCGAGACCTGACTGAGGCTTCTTGGGAGAGATGAG 840
QY 765 AAGCCAAAGGGGAACCTTGGAGAGCTTCAATGATGAGAACCTTGGGAGATGAGTGGCT 824
DB 841 AAGCCAAAGGGGAACCTTGGAGAGCTTCAATGATGAGAACCTTGGGAGATGAGTGGCT 900
QY 825 GACCTGTTCTCTGCGGGATGCTGACCACTCGACCACTGAGCTGAGCTGAGGCTTCTGCTC 884
DB 901 GACCTGTTCTCTGCGGGATGCTGACCACTCGACCACTGAGCTGAGCTGAGGCTTCTGCTC 960
QY 885 ATGATCTCTACATCCGATGTGACGCGCGTGTCCAAACAGGAGATGAGAGCTGATAGGG 944
DB 961 ATGATCTCTACATCCGATGTGACGCGCGTGTCCAAACAGGAGATGAGAGCTGATAGGG 1020
QY 945 CAGTGGCGGACACAGAGATGGGTGACAGGCTCACATGCCCTACACCACTGCCGTGATT 1004
DB 1021 CAGTGGCGGACACAGAGATGGGTGACAGGCTCACATGCCCTACACCACTGCCGTGATT 1080
QY 1005 CATGAGGTGACGCGCTTTGGGACATCGTCCCTGGTGTGACCCATATGACATCCCGT 1064
DB 1081 CATGAGGTGACGCGCTTTGGGACATCGTCCCTGGTGTGACCCATATGACATCCCGT 1140
QY 1065 GACATCGAAGTACAGGGCTTCCGATCCCTAAGGGAACGACATCATCAACCTGTCA 1124
DB 1141 GACATCGAAGTACAGGGCTTCCGATCCCTAAGGGAACGACATCATCAACCTGTCA 1200
QY 1125 TCGGTGCTGAAGATGAGGCGCTCTGGGAGAGGCGCTTCCGCTTCCACCCCGAACAATTC 1184
DB 1201 TCGGTGCTGAAGATGAGGCGCTCTGGGAGAGGCGCTTCCGCTTCCACCCCGAACAATTC 1260
QY 1185 CTGATGCCACAGGCGCACTTTGTGAAGCGGAGGCGCTTCCGCTTCCACAGAGGCGCG 1244
DB 1261 CTGATGCCACAGGCGCACTTTGTGAAGCGGAGGCGCTTCCGCTTCCACAGAGGCGCG 1320
QY 1245 CGTGCATGCTCGGGAGCGCTTGGCGCGCATGAGCTTTCCTTCTTCACTTCCCTG 1304
DB 1321 CGTGCATGCTCGGGAGCGCTTGGCGCGCATGAGCTTTCCTTCTTCACTTCCCTG 1380
QY 1305 CTGAGCACTTCACTGCTTGGTGGCGCACTGGAGAGCGCGCGCGCGAGCCACCATGCTGTC 1364
DB 1381 CTGAGCACTTCACTGCTTGGTGGCGCACTGGAGAGCGCGCGCGCGAGCCACCATGCTGTC 1440
QY 1365 TTTGCTTTCTGCTGAGCGCCATCCCTATGAGCTTTGTGCTGCTGCGCGCGCTAGAAATGG 1424
DB 1441 TTTGCTTTCTGCTGAGCGCCATCCCTATGAGCTTTGTGCTGCTGCGCGCGCTAGAAATGG 1500
QY 1425 GTACCTAGTCCCGAGCTGCTCCTAGCCAGAGCTCTTAATGTAC 1469
DB 1501 GTACCTAGTCCCGAGCTGCTCCTAGCCAGAGCTCTTAATGTAC 1545

RESULT 12

AAQ87730

ID AAQ87730 standard; cDNA; 1494 BP.

XX AAQ87730;

AC AAQ87730;

XX 25-MAR-2003 (revised)

DT 15-NOV-1995 (first entry)

XX Human auxillary cytochrome P450 species 2D6 variant 1 coding region.

XX Human cytochrome P450; amplification; PCR; primer; expression vector;

KW yeast NADPH-P450 reductase; safety; fusion protein; metabolite;

KW carcinogen; mutagen; liver metabolism; ds.

XX Homo sapiens.

OS

XX FH Key Location/Qualifiers
FT variation 886
FT /tag= a
FT /note= "T to C change in variant 1 changes amino acid
FT from Cys to Arg"
XX
XX
XX EP644267-A2.
XX
XX 22-MAR-1995.
XX
XX 20-JUL-1994; 94EP-00111298.
XX
XX 20-JUL-1993; 93JP-00201120.
XX PR 21-JUL-1993; 93JP-00180246.
XX PR 30-JUL-1993; 93JP-00208279.
XX
XX (SUMO) SUMITOMO CHEM CO LTD.
XX
XX Hayaashi K, Sakaki T, Yabueaki Y, Komai K, Kaneko H, Nakatsuka I;
XX
XX WPI; 1995-116991/16.
XX DR P-PSDB; AAR72376.
XX
XX Evaluation of safety of a chemical cpd. - using recombinant yeast
PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase.
XX
XX Example; Page 87-89; 124pp; English.
XX
XX The nucleotide sequence of the cDNA coding region for the human auxiliary
CC cytochrome P450 species 2D6 variant 1. The gene contains a change at base
CC 886 from T to C as compared to the wild type sequence (AA087729). This
CC changes the amino acid residue 296 from Cys to Arg. The cDNA was
CC amplified by PCR using the primers AA087763-6. The product was cloned
CC into the yeast expression vectors pAH5N or pAHR to produce the vectors
CC p2D6 variant 1 for the expression of the cytochrome P450 alone or p2D6R
CC variant 1 for co-expression with the yeast NADPH-P450 reductase. The
CC vectors are used in a method for evaluating the safety of a chemical
CC compound by reacting the chemical compound with recombinantly produced
CC human cytochrome P450 molecular species 1A2 (AA087714), 2C9 (AA087715),
CC 2E1 (AA087716), or 3A4 (AA087717) or their auxiliary species and variants
CC (AA087718-32), and yeast NADPH-P450 reductase, either as a fused protein
CC or in cell extracts, and analysing the resulting metabolite to assess the
CC safety of the chemical compound. The method is useful for determining
CC whether the chemical compound, or its metabolite, will be converted into
CC a carcinogenic or mutagenic form through metabolism in the liver.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 1494 BP; 248 A; 509 C; 446 G; 291 T; 0 U; 0 Other;
SQ

Query Match 76.3%; Score 1173.2; DB 2; Length 1494;
Best Local Similarity 89.6%; Pred. No. 1.4e-213;
Matches 1338; Conservative 0; Mismatches 3; Indels 153; Gaps 1;
78 ATGGGGCTAGAGCACTGGTCCCTGGCCGCTGATAGTGGCCATCTTCTGCTCTCTGGTG 137
1 ATGGGGCTAGAGCACTGGTCCCTGGCCGCTGATAGTGGCCATCTTCTGCTCTCTGGTG 60
138 GACCTGATGACACGGGGCAACGCTGGCTGACGCTACTACACAGGCCCTGCACTG 197
61 GACCTGATGACACGGGGCAACGCTGGCTGACGCTACTACACAGGCCCTGCACTG 120
198 CCGGGCTGGGCAACCTGCTGATGCTGACCTTCCAGACACACACTACTGCTTCCAGCAG 257
121 CCGGGCTGGGCAACCTGCTGATGCTGACCTTCCAGACACACACTACTGCTTCCAGCAG 180
258 TTGCGCGCGCTTCCGGGACGCTTTCAGCTGACGCTGGCTGGACGCGCGTGGTCTG 317
181 TTGCGCGCGCTTCCGGGACGCTTTCAGCTGACGCTGGCTGGACGCGCGTGGTCTG 240
318 CTCATATGGCTGGCGGCTGGCGGAGCGCTGGTACCCACGGCGAGACACCGCCGAC 377
241 CTCATATGGCTGGCGGCTGGCGGAGCGCTGGTGAACCCACGGCGAGGACACCGCCGAC 300

QY 378 CGCCCGCCTGTGCCATCACCAGATCCCTGGGTTTTGGGCCCGCTTCCCAA----- 428
DB 301 CGCCCGCCTGTGCCATCACCAGATCCCTGGGTTTTGGGCCCGCTTCCCAAAGGGGTGTTTC 360
QY 429 ----- 428
DB 361 CTGGCGCGCTATATGGGCCCGCTGGCGGAGCAGAGCGCGCTTCTCCGTCTCCACCTTGGCG 420
QY 429 ----- 428
DB 421 AACTTGGGCGCTGGGCAAGAAGTCGTGGAGCAGTGGTGACCGAGAGGCGCGCTCCCTT 480
QY 429 -----GGACGCCCCCTTTCGCCCCCAACCGGTCTCTTGGACAAA 464
DB 481 TGTGGCGCTTTCGCCAAACCACTCCGAGCGCCCCCTTTCGCCCCCAACCGGTCTCTTGGACAAA 540
QY 465 GCGGTGAGCAACGCTGATCGCTCCCTCACTCGCGGCGCGCTTCGAGTACGACGCCCT 524
DB 541 GCGGTGAGCAACGCTGATCGCTCCCTCACTCGCGGCGCGCTTCGAAATACGACGACCT 600
QY 525 GCGTTCCTCAGGCTGCTGACCTAGCTCAGGAGGACTGAAGGAGGAGTGGGCTTTCTG 584
DB 601 GCGTTCCTCAGGCTGCTGACCTAGCTCAGGAGGAGTGAAGGAGGAGTGGGCTTTCTG 660
QY 585 GCGGAGGTGCTGAATGCTGCTCCCTGCTCCATATCCAGCGCTGGCTGGCAAGTTC 644
DB 661 GCGGAGGTGCTGAATGCTGCTCCCTGCTCCATATCCAGCGCTGGCTGGCAAGTTC 720
QY 645 CTACGCTTCCAAAAGGCTTTCTGACCCAGCTGGATGAGCTGCTAACTGAGACAGGATG 704
DB 721 CTACGCTTCCAAAAGGCTTTCTGACCCAGCTGGATGAGCTGCTAACTGAGACAGGATG 780
QY 705 ACCTGGGACCCAGCCCGGAGCTGCTGAGCGCTTCTTGGCAGAGATGGAG 764
DB 781 ACCTGGGACCCAGCCCGGAGCTGCTGAGCGCTTCTTGGCAGAGATGGAG 840
QY 765 AAGGCCAAGGGGAACCTTGAGAGAGCTTCAATGATGAGACCTCGCATGCTGGTGGCT 824
DB 841 AAGGCCAAGGGGAACCTTGAGAGAGCTTCAATGATGAGACCTCGCATGCTGGTGGCT 900
QY 825 GACCTGTTCTCTGCGGGGATGGTACCACTCGACCGCTGGCTGGGCGCTCTCTGCTC 884
DB 901 GACCTGTTCTCTGCGGGGATGGTACCACTCGACCGCTGGCTGGGCGCTCTCTGCTC 960
QY 885 ATGATCTTACATCCGAGTGGCGCGCTGTCACACAGGAGATCGACGCTGATAGGG 944
DB 961 ATGATCTTACATCCGAGTGGCGCGCTGTCACACAGGAGATCGACGCTGATAGGG 1020
QY 945 CAGGTGGCGGACACAGAGATGGGTGACCGAGCTCAGATGCCCTACACCTGCGCGTATT 1004
DB 1021 CAGGTGGCGGACACAGAGATGGGTGACCGAGCTCAGATGCCCTACACCTGCGCGTATT 1080
QY 1005 CATGAGTGGCAGCGCTTTGGGGGACATCGTCCCTCGGCTGTGACCCATATGACATCCCGT 1064
DB 1081 CATGAGTGGCAGCGCTTTGGGGGACATCGTCCCTCGGCTGTGACCCATATGACATCCCGT 1140
QY 1065 GACATCGAAGTACAGGGCTTCCGATCCCTAAGGGAACGACACTCATACCAACTGTCA 1124
DB 1141 GACATCGAAGTACAGGGCTTCCGATCCCTAAGGGAACGACACTCATACCAACTGTCA 1200
QY 1125 TCGGTGCTGAAGGATGAGCGCTTGGGAGAGCGCTTCCGCTTCCACCCGGAACACTTC 1184
DB 1201 TCGGTGCTGAAGGATGAGCGCTTGGGAGAGCGCTTCCGCTTCCACCCGGAACACTTC 1260
QY 1185 CTGATGCCAGGCGCACTTTGTGAAGCGGAGGCGCTTCTCGCTTCTCAGCAGGCGCG 1244
DB 1261 CTGATGCCAGGCGCACTTTGTGAAGCGGAGGCGCTTCTCGCTTCTCAGCAGGCGCG 1320
QY 1245 CGTGATGCCCTCGGGGAGCGCTTGGCGCGCATGAGAGCTTCTCTCTTCTTCACTCCCTG 1304
DB 1321 CGTGATGCCCTCGGGGAGCGCTTGGCGCGCATGAGAGCTTCTCTCTTCTTCACTCCCTG 1380

QY 1305 CTGACGACATTCAGCTTCCTGCTGCCACTGACAGACCCCGCCAGCCACCATGGTGTG 1364
DB |||||
DB 1381 CTGACGACATTCAGCTTCCTGCTGCCACTGACAGACCCCGCCAGCCACCATGGTGTG 1440
QY 1365 TTGCTTTCTGCTGGTGAACCCATCCCTCATGAGCTTTGTGTGCTGCCCGGTAG 1418
DB |||||
DB 1441 TTGCTTTCTGCTGGTGAACCCATCCCTCATGAGCTTTGTGTGCTGCCCGGTAG 1494

RESULT 13

AAT28396
ID AAT28396 standard; DNA; 1494 BP.

XX
AC AAT28396;

XX
DT 11-OCT-1996 (first entry)

XX
DE Human cytochrome P450 molecular species 2D6 variant #2 gene.

XX
KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;
KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;
KW evaluation; safety; fusion protein; metabolite; detoxification;
KW carcinogenic; ds.

XX
OS Homo sapiens.

XX
PN JP08056695-A.

XX
PD 05-MAR-1996.

XX
PF 15-JUL-1994; 94JP-00164184.

XX
PR 20-JUL-1993; 93JP-00201120.

PR 30-JUL-1993; 93JP-00208279.

PR 17-JUN-1994; 94JP-00136053.

XX
PA (SUMO) SUMITOMO CHEM CO LTD.

XX
WPI; 1996-182311/19.

DR
P-PSDB; AAR93183.

XX
Novel method for the evaluation of the safety of a cpd. - using a human

PT cytochrome P450 and yeast NADPH reductase to determine whether the

PT analyte cpd. is detoxified or metabolised to a carcinogen.

XX
Example 1; Page 51-53; 74pp; Japanese.

XX
This is the nucleotide sequence of the human cytochrome P450 molecular
species 2D6 variant #2 which encodes a protein of 497 amino acids. The
gene was amplified from a human liver derived cDNA library as 2 fragments
of 0.4 and 0.9 kb using primers AAT28381-1 and AAT28382-1. The prod. was cloned into the
yeast expression vector pAAH5N to generate plasmid p2D6 for prodn. of the
cytochrome only or into the vector pAHR to generate the plasmid p2C6R
for co-prodn. with the yeast NADPH-P450 reductase. The sequence is placed
under control of the yeast ADH gene promoter and terminator. The vectors
are used in a method for evaluating the safety of a cpd. by reacting the
test cpd. with recombinantly produced human cytochrome P450 mol. species
1A2 (AAT28380), 2C9 (AAT28381), 2E1 (AAT28382), 3A4 (AAT28383) or their
variants (AAT28384-98) together with yeast NADPH-P450 reductase (either
as a fused protein or as a cell extract) and analysing the resultant
metabolite. The cpd. is considered "safe" if it is detoxified or not
rendered carcinogenic or "unsafe" if it is not detoxified or is
metabolised to a carcinogenic cpd

XX
Sequence 1494 BP; 248 A; 509 C; 446 G; 291 T; 0 U; 0 Other;

Query Match 76.3%; Score 1173.2; DB 2; Length 1494;

Best Local Similarity 89.6%; Pred. No. 14e-213;

Matches 1338; Conservative 0; Mismatches 23; Indels 153; Gaps 1;

QY 78 ATGGGGCTAGAGCACTGGTGGCCCTGGCCCGGTGATAGTGGCCATCTTCTGCTCTGGTG 137

DB 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCCCGGTGATAGTGGCCATCTTCTGCTCTGGTG 60

QY 138 GACCTGATGACACCGCGCCCAACGCTGGCTGACGCTACTCACCAGGCCCCCTGCCACTG 197
DB |||||
DB 61 GACCTGATGACACCGCGCCCAACGCTGGCTGACGCTACTCACCAGGCCCCCTGCCACTG 120
QY 198 CCGGGCTGGGCAACCTGCTGCAATGTGGACTTCAGAAACACACCATATCTGCTTCGACG 257
DB |||||
DB 121 CCGGGCTGGGCAACCTGCTGCAATGTGGACTTCAGAAACACACCATATCTGCTTCGACG 180
QY 258 TTGCGGCGCGCTTTGGGGGAGGTGTTGAGCTGCGAGCTGGCTTGGAGCGCGGTGGTGTG 317
DB |||||
DB 181 TTGCGGCGCGCTTTGGGGGAGGTGTTGAGCTGCGAGCTGGCTTGGAGCGCGGTGGTGTG 240
QY 318 CTCAATGGGCTGGGCGCGCTGGCGAGCGCTGTGTGACCCACGCGGAGGACACCGCGGAC 377
DB |||||
DB 241 CTCAATGGGCTGGGCGCGCTGGCGAGCGCTGTGTGACCCACGCGGAGGACACCGCGGAC 300
QY 378 CGCCCGCTGTGCCCCATCACCCAGATCCTGGGTTTTGGGCGCGCTTCCCAA----- 428
DB |||||
DB 301 CGCCCGCTGTGCCCCATCACCCAGATCCTGGGTTTTGGGCGCGCTTCCCAAAGGGGTGTTTC 360
QY 429 ----- 428
DB 361 CTGGCGCGCTATGGGCGCGCTGGCGAGCAGAGCGCTTCTCCGTTCTCCACTTGGCGC 420
QY 429 ----- 428
DB 421 AACTTGGGCTGGGCAAGAGTGGCTGGAGCAGATGGGTTGACCGAGAGGCGCGCTGCCCTT 480
QY 429 -----GGACGCGCGCTTTCCGCCCAACGCTTCTTGGACAAA 464
DB 481 TGTGCGCGCTTTCGCAACCACTCGGAGCGCGCTTTCCGCCCAACGCTTCTTGGACAAA 540
QY 465 GCGTGGAGCAACGTGATCGCTTCCCTCACTGGGCGCGCGCTTTCGAGTAGCAGACCCCT 524
DB |||||
DB 541 GCGTGGAGCAACGTGATCGCTTCCCTCACTGGGCGCGCGCTTTCGAAATACGACGACCT 600
QY 525 GCGTTCCTCAGGCTGGGACCTAGCTCAGAGGAGCTGAAGAGGAGTGGGCTTCTGTG 584
DB |||||
DB 601 CGTTCTCAGGCTGGGAGCTTAGCTCAGAGGAGGAGTGAAGAGGAGTGGGCTTCTGTG 660
QY 585 GCGAGGTGCTGAATGTGTCTCCCGTCTCTGATATCCAGCGCTGGCTGGCAAGTGC 644
DB |||||
DB 661 GCGAGGTGCTGAATGTGTCTCCCGTCTCTGATATCCAGCGCTGGCTGGCAAGTGC 720
QY 645 CTACGCTTCCAAAGGCTTCTCTGACCGAGCTGAGCTGCTAACTGACGACAGGATG 704
DB |||||
DB 721 CTACGCTTCCAAAGGCTTCTCTGACCGAGCTGAGCTGCTAACTGACGACAGGATG 780
QY 705 ACCTGGGACCCAGCCAGCCCGGAGACCTGACTGAGGCTTCTTGGGAGAGATGGAG 764
DB |||||
DB 781 ACCTGGGACCCAGCCAGCCCGGAGACCTGACTGAGGCTTCTTGGGAGAGATGGAG 840
QY 765 AAGGCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACTTGCGCATAGTGGTGCCT 824
DB |||||
DB 841 AAGGCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACTTGCGCATAGTGGTGCCT 900
QY 825 GACCTGTTCTTGGCGGAGTGGTGAACCACTCGACCGCTGGCTGGGCGCTCTGCTGC 884
DB |||||
DB 901 GACCTGTTCTTGGCGGAGTGGTGAACCACTCGACCGCTGGGCGCTCTGCTGC 960
QY 885 ATGATCTTACATCCGAGTGGACGCGCTGTGTCAACAGGAGATCGACGCTGATAGGG 944
DB |||||
DB 961 ATGATCTTACATCCGAGTGGACGCGCTGTGTCAACAGGAGATCGACGCTGATAGGG 1020
QY 945 CAGTGGCGGACACAGAGATGGGTGACGAGCTCACATGCCCTACACACTGCCGTGATT 1004
DB |||||
DB 1021 CAGTGGCGGACACAGAGATGGGTGACGAGCTCACATGCCCTACACACTGCCGTGATT 1080
QY 1005 CATGAGTGCAGCGCTTTGGGACATGTCCTCCCTGGGTGTGACCCATATGACATCCCGT 1064
DB |||||
DB 1081 CATGAGTGCAGCGCTTTGGGACATGTCCTCCCTGGGTGTGACCCATATGACATCCCGT 1140


```
QY 1065 GACATCGAAGTACAGGGCTTCGCATCCCTAAGGGAACGACACTCATCAACCACTGTCA 1124
Db 1141 GACATCGAAGTACAGGGCTTCGCATCCCTAAGGGAACGACACTCATCAACCACTGTCA 1200
QY 1125 TCGGTCTCAAGATCAGAGCGGCTCTGGAGAGCCCTTCGGCTTCCACCCCGAACAACCTTC 1184
Db 1201 TCGGTCTCAAGATCAGAGCGGCTCTGGAGAGCCCTTCGGCTTCCACCCCGAACAACCTTC 1260
QY 1185 CTGGATGCCAGGCGCATTGTGAAAGCCGAGGCTTCTGCTCTTCTCAGAGGCGCG 1244
Db 1261 CTGGATGCCAGGCGCATTGTGAAAGCCGAGGCTTCTGCTCTTCTCAGAGGCGCG 1320
QY 1245 CGTGATGCTCTGGAGAGCCCTTCGGCCGATGAGCTTCTCTTCTTCACTCCCTG 1304
Db 1321 CGTGATGCTCTGGAGAGCCCTTCGGCCGATGAGCTTCTCTTCTTCACTCCCTG 1380
QY 1305 CTGCAGCACTTCACTTCTCGGTGCGCCACTGGACAGCCCGCCGACCCACCATGCTGTC 1364
Db 1381 CTGCAGCACTTCACTTCTCGGTGCGCCACTGGACAGCCCGCCGACCCACCATGCTGTC 1440
QY 1365 TTGCTTTCTGCTGATGACCCCATCCCTATGAGCTTTGCTGTGCTGCTGCTGCTAG 1418
Db 1441 TTGCTTTCTGCTGATGACCCCATCCCTATGAGCTTTGCTGTGCTGCTGCTGCTAG 1494

RESULT 14
ABQ72216
ID ABQ72216 standard; cDNA; 1494 BP.
XX
AC ABQ72216;
XX
DT 02-SEP-2002 (first entry)
DE Human CYP2D6 gene coding sequence, SEQ ID NO:2.
XX
KW Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;
KW Chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;
KW antiarrhythmic; arrhythmia; adrenoreceptor antagonist; hypertension;
KW tricyclic antidepressant; procainamide; drug induced lupus syndrome;
KW environmentally linked disease; Parkinson's disease; haplotyping;
KW genotyping; haplotype; genetic variant; single nucleotide polymorphism;
KW SNP; drug screening; drug discovery; gene; ss.
XX
OS Homo sapiens.
XX
FH Key
FH CDS
FT 1..1494
FT /*tag= a
FT /product= "CYP2D6"
FT replace(19, A)
FT /*tag= b
FT /label= PS7
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V7M"
FT replace(31, A)
FT /*tag= c
FT /label= PS8
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V11M"
FT replace(100, T)
FT /*tag= d
FT /label= PS9
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution P34S"
FT replace(263, A)
FT /*tag= e
FT /label= PS12
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution R88H"
FT replace(271, A)
FT /*tag= f
FT /label= PS13
FT /note= "Known single nucleotide polymorphism (SNP);
```

```
FT causes the amino acid substitution L91M"
FT replace(281, G)
FT /*tag= g
FT /label= PS14
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution H94R"
FT replace(294, G)
FT /*tag= h
FT /label= PS15
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace(311, C)
FT /*tag= i
FT /label= PS16
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V104A"
FT replace(319, T)
FT /*tag= j
FT /label= PS17
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS18 causes the amino acid substitution
FT T107F"
FT replace(320, T)
FT /*tag= k
FT /label= PS18
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS17 causes the amino acid substitution
FT replace(325, G)
FT /*tag= l
FT /label= PS19
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution I109V"
FT replace(333, C)
FT /*tag= m
FT /label= PS20
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace(336, T)
FT /*tag= n
FT /label= PS21
FT /note= "Known single nucleotide polymorphism (SNP)"
FT replace(358, A)
FT /*tag= o
FT /label= PS27
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution F120I"
FT replace(382, C)
FT /*tag= p
FT /label= PS28
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution W128R"
FT replace(406, A)
FT /*tag= q
FT /label= PS29
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS30 causes the amino acid substitution
FT V136I"
FT replace(408, C)
FT /*tag= r
FT /label= PS30
FT /note= "Known single nucleotide polymorphism (SNP);
FT together with PS29 causes the amino acid substitution
FT V136I"
FT replace(451, G)
FT /*tag= s
FT /label= PS31
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution Q151E"
FT replace(463, A)
FT /*tag= t
FT /label= PS32
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution E155K"
FT replace(696, C)
FT variation
```


Db 421 AACTTGGCCCTGGCAAGATGCTGGAGCAGTGGGTGACCGAGAGGCCGCTGCCTT 480
QY 429 -----GGACGCCCTTTTGGCCCCAACGGTCTCTTGGACAAA 464
Db 481 TGTGCGGCTTGGCCAACCACTCGGAGCGCCCTTTGCCCCAACGGTCTCTTGGACAAA 540
QY 465 GCGGTGAGCAAGTGAATCGCTCCCTACCTGCGGGCGCGCTTCGAGTACGACGCCCT 524
Db 541 GCGGTGAGCAAGTGAATCGCTCCCTACCTGCGGGCGCGCTTCGAGTACGACGCCCT 600
QY 525 GCTTTCCTCAGGCTGCTGGAATAGCTCAGAGGGAAGTGAAGGAGGAGTGGGCTTTCTG 584
Db 601 GCTTTCCTCAGGCTGCTGGAATAGCTCAGAGGGAAGTGAAGGAGGAGTGGGCTTTCTG 660
QY 585 GCGAGGTGCTGAATGCTGCTCCGCTCTCTGATATCCAGCGCTGGCTGGCAAGTTC 644
Db 661 GCGAGGTGCTGAATGCTGCTCCGCTCTCTGATATCCAGCGCTGGCTGGCAAGTTC 720
QY 645 CTACGCTTCCAAAAGGCTTTCTGTACCCAGCTGGATGAGCTGCTAACTGAGCAGAGATG 704
Db 721 CTACGCTTCCAAAAGGCTTTCTGTACCCAGCTGGATGAGCTGCTAACTGAGCAGAGATG 780
QY 705 ACCTGGGACCCAGCCACGCCCCCGAGACCTGACTGAGGCTTCTGCGCAGAGATGGAG 764
Db 781 ACCTGGGACCCAGCCACGCCCCCGAGACCTGACTGAGGCTTCTGCGCAGAGATGGAG 840
QY 765 AAGGCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGTGGCT 824
Db 841 AAGGCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGTGGCT 900
QY 825 GACCTGTTCCTGCGGATGGTGACCACTCGACACGCTGGCTGGGCGCTCCTGCTC 884
Db 901 GACCTGTTCCTGCGGATGGTGACCACTCGACACGCTGGCTGGGCGCTCCTGCTC 960
QY 885 ATGATCCTACATCGGATGGTGAGCGCGGTGTCAAACAGAGATCGACGATGATAGGG 944
Db 961 ATGATCCTACATCGGATGGTGAGCGCGGTGTCAAACAGAGATCGACGATGATAGGG 1020
QY 945 CAGGTGCGGCGACAGAGATGGGTGACAGGCTCACATGCCCTACACCTGCCGTGATT 1004
Db 1021 CAGGTGCGGCGACAGAGATGGGTGACAGGCTCACATGCCCTACACCTGCCGTGATT 1080
QY 1005 CATGAGGTGACAGCGCTTTTGGGACATCGTCCCTCGGTGTGACCCATATGACATCCCGT 1064
Db 1081 CATGAGGTGACAGCGCTTTTGGGACATCGTCCCTCGGTGTGACCCATATGACATCCCGT 1140
QY 1065 GACATCGAAGTACAGGCTTCGCAATCCCTAAGGGAACGACATCATCAACCAACCTGTCA 1124
Db 1141 GACATCGAAGTACAGGCTTCGCAATCCCTAAGGGAACGACATCATCAACCAACCTGTCA 1200
QY 1125 TCGGTGCTGAGGATGAGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCCGACACTTC 1184
Db 1201 TCGGTGCTGAGGATGAGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCCGACACTTC 1260
QY 1185 CTGGATGCCAGGGCCACTTTGTGAAGCGGAGGCGCTTCTGCGCTTTCTCAGCAGGCGCG 1244
Db 1261 CTGGATGCCAGGGCCACTTTGTGAAGCGGAGGCGCTTCTGCGCTTTCTCAGCAGGCGCG 1320
QY 1245 CGTGATGCCCTGGGAGGCCCTGCGCGCATGGAGCTTTCCTCTTTCTTCACTCCCTG 1304
Db 1321 CGTGATGCCCTGGGAGGCCCTGCGCGCATGGAGCTTTCCTCTTTCTTCACTCCCTG 1380
QY 1305 CTGAGCACTTCAGCTTCTCGGTGCCACTGGAGAGCCCGCGCCAGCCACCATGGTGTTC 1364
Db 1381 CTGAGCACTTCAGCTTCTCGGTGCCACTGGAGAGCCCGCGCCAGCCACCATGGTGTTC 1440
QY 1365 TTGTGCTTTCTGCTGACCCCTATCCCTATGAGCTTTGTGCTGTGCCCGGTAG 1418
Db 1441 TTGTGCTTTCTGCTGAGCCCATCCCTATGAGCTTTGTGCTGTGCCCGGTAG 1494

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 25, 2004, 02:15:53 ; Search time 70.8984 Seconds
(without alignments)
13680.175 Million cell updates/sec

Title: US-09-820-788A-1
Perfect score: 2886
Sequence: 1 cctgcctgctctctgtgcc.....aaaaaaaaaaaaaaaaaaaaa 1537

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US09820788/runat.24022004.141426.20037/app_query.fasta_1.12174
-DB=SPTRMBL_25 -QFWT=fastan -SURFIX=rspt -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09820788 @CNG 1 1 785 @runat.24022004.141426.20037 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
SPTRMBL_25:
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES
Result No. Score Query Match Length DB ID Description
-----
1 2283 79.1 500 4 Q16753 Q16753 homo sapien

```

```

2 2197.5 76.1 497 6 Q8WNRS
3 2138.5 74.1 497 6 Q865W1
4 1769 61.3 500 6 Q29454
5 1751 60.7 500 6 Q9TUJ4
6 1719.5 59.6 373 4 Q16804
7 1711 59.3 500 11 Q9JKY7
8 1705 59.1 500 11 Q9IWB7
9 1697 58.8 500 6 Q9TJUS
10 1633.5 56.6 500 11 Q8VCX0
11 1625 56.3 500 11 Q8CIM7
12 1602.5 55.2 500 11 Q9DBJ5
13 1622 56.2 500 11 Q9DBJ5
14 1582.5 55.5 504 11 Q921V1
15 1582.5 54.8 504 11 Q8BVD2
16 1567.5 54.3 504 11 Q64530
17 1157.5 40.1 505 13 Q7SYW2
18 932.5 32.3 497 13 Q7ZU60
19 927.5 32.1 498 13 Q9PTR2
20 921 31.9 496 13 Q803J0
21 881.5 30.5 498 13 Q9IAX8
22 878.5 30.4 498 13 Q9PTR1
23 864 29.9 492 13 Q98T96
24 864 29.9 492 13 Q7ZV97
25 857.5 29.7 497 13 Q9FV10
26 845 29.3 501 11 Q8CC91
27 845 29.3 501 11 Q8BR78
28 836.5 29.0 502 4 Q8TF13
29 834 28.9 501 11 Q9QXF7
30 830.5 28.8 505 13 Q9IAT1
31 828 28.7 502 11 Q924D1
32 824.5 28.6 491 6 Q29516
33 822.5 28.5 503 13 Q802X8
34 821.5 28.5 491 11 Q64460
35 821.5 28.5 497 13 Q9PVH9
36 820 28.4 529 11 Q9CRQ9
37 818.5 28.4 503 13 Q800W8
38 817.5 28.3 491 11 Q64463
39 810.5 28.1 493 6 Q8SQ67
40 807.5 28.0 491 6 Q29532
41 806.5 27.9 491 11 Q9WUD0
42 798.5 27.7 505 13 Q7SXX7
43 798.5 27.7 506 13 Q9FV11
44 798.5 27.7 521 13 Q802U6
45 794.5 27.5 490 4 Q8WMB1

```

ALIGNMENTS

```

RESULT 1
Q16753 PRELIMINARY; PRT; 500 AA.
AC Q16753;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE Debrisoquine 4-hydroxylase mutant allele (CYP2D6-MAL)
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gonzalez F.J.
RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; M33189; AAA35737.1; -
DR HSSP; P00179; 1D76.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR pfam; PF00067; p450; 1.

```

DR PRINTS; PR01686; EP450ICYP2D.

DR PRINTS; PR0385; P450.

DR PROSITE; PS00086; CYTOCHROME P450; 1.

KW Heme; Monooxygenase; Oxidoreductase.

SQ SEQUENCE 500 AA; 55889 MW; D5293B9BF74692C8 CRC64;

Alignment Scores:

Pred. No.: 3,6e-148 Length: 500
Score: 2283.00 Matches: 444
Percent Similarity: 89.0% Conservative: 1
Best Local Similarity: 88.80% Mismatches: 1
Query Match: 79.11% Indels: 54
Gaps: 1

US-09-820-788a-1 (1-1537) x Q16753 (1-500)

QY 78 ATGGGGCTAGAGCACTGGTCCCTCGCCCTGATAGTGGCCATCTTCTCTCTCTGCTG 137
Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
QY 138 GACCTGATGACCGCGCCAAAGCTGGGCTGCAGCTACTACACAGGCCCTCCCTGCTG 197
Db 21 AspLeuMetHisArgArgGlnArgTTPAlaAlaArgTyrSerProGlyProLeu 40
QY 198 CCCGGCTGGGCAACTCTGCTGATGGACTTCCAGAACACACCATACTGCTTCGACG 257
Db 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
QY 258 TTGGCGCCCGCTTCCGGGAGCTGTTCAGCTGAGCTGGCTGGAGCCCGCTGGTGGT 317
Db 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTTPProValVal 80
QY 318 CTCATGGCTGGCGCGCTGGGAGCGCTGTGACCCAGCGGAGGACACCGCGAC 377
Db 81 LeuAsnGlyLeuAlaAlaValArgGluAlaMetValThrArgGlyGluAspThrAlaAsp 100
QY 378 CGCCCGCTGTGCCATCACCCAGATCCTGGGTTTGGCGCGCTTCCCAAGGA- 431
Db 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGly 120
QY 431 ----- 431
Db 121 ArgGlyValProGlyAlaLeuTTPAlaArgValAlaArgAlaGluAlaLeuArgLeu 140
QY 431 ----- 431
Db 141 HisLeuAlaGlnLeuGlyProGlyGlnGluValAlaGlyAlaValGlyAspArgGly 160
QY 432 -----CGCCCTTTCGCCCTTCCGCCCAACGGTCTC 455
Db 161 ArgLeuProLeuCysArgLeuArgGlnProLeuArgArgProPheArgProAsnGlyLeu 180
QY 456 TTGACAAAGCGTGAGCAAGCTGATCCCTCCCTCACCTCGGGCGCGCTTCAGTAC 515
Db 181 LeuAspGlyAlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyr 200
QY 516 GACGACCTCTCTCTCTCGCTGCTGAGCTAGCTCAGGAGGAGTGAAGAGAGTGG 575
Db 201 AspAspProArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSer 220
QY 576 GGCTTTCGCGGAGTGTGAATGCTGCTCCCTCCCTCTGATATCCAGCGCTGGCT 635
Db 221 GlyPheLeuArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAla 240
QY 636 GGCAAGGTCTACGCTTCCAAAGGCTTCTGTGACCCAGCTGGATGAGTCTTAACGTAG 695
Db 241 GlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGlu 260
QY 696 CACAGGATGACTGGAGCCAGCCAGCCCGGAGACCTGATGAGGCGCTTCTGGCA 755
Db 261 HisArgMetThrTTPAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAla 280
QY 756 GAGATGGAGAGGCCAGGGGAACCTCGAGAGCAGCTTCATGATGAGAACCTCGGCATA 815

Db 281 GluMetGluLysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIle 300
QY 816 GTGTGGCTGACCTGTTCTCTGCGGATGGTGACCACTCGACCGCTGGCTGGCGGC 875
Db 301 ValValAlaAspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTTPGly 320
QY 876 CTCTGTCTGATGCTTACATCCGATCGGATGTCAGCGCTGTCACAGGAGATCGACGAC 935
Db 321 LeuLeuLeuMetIleLeuHisProAspValGlnArgValGlnGlnGluIleAspAsp 340
QY 936 GTGATAGGCGAGGTGCGCGCACACAGATGGTGACCGAGCTCACATGCTCCCTACCACT 995
Db 341 ValIleGlyGlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThr 360
QY 996 GCGTGTATCATGAGTGCAGCGCTTTGGGAGCATCTCCCTCCCTGGGTGTGACCCATATG 1055
Db 361 AlaValIleHisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMet 380
QY 1056 ACATCCCTGACATCGAAGTACAGGCTTCGGATCCCTAAGGAAACGACACTCATCACC 1115
Db 381 ThrSerArgAspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThr 400
QY 1116 AACCTGTCTCATCGTGTGAAGGATGAGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCC 1175
Db 401 AsnLeuSerValLeuLysAspGluAlaValTTPGluLysProPheArgPheHisPro 420
QY 1176 GAACACTTCTGGATGCCAGGCGCACTTTGTGAAGCCGAGGCGCTTCTGCTCTCTCA 1235
Db 421 GluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSer 440
QY 1236 GCAGGCGCGTGCATGCTCGGGAGCCCTCGCGGAGCCCTCGCGGATGAGCTTCTCTTCTC 1295
Db 441 AlaGlyArgArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePhe 460
QY 1296 ACCTCCCTGTGCGAGCACTTTCAGTCTCGGTGCCACTGCAGAGCCCGCGCCAGCCAC 1355
Db 461 ThrSerLeuLeuGlnHisPheSerValProThrGlyGlnProArgProSerHis 480
QY 1356 CATGGTCTTCTGCTTCTGTCAGCCCACTCCCTATCAGCTTTGTGTGTCGCCCGC 1415
Db 481 HisGlyValPheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 500

RESULT 2
Q8WNR5 Q8WNR5 PRELIMINARY; PRT; 497 AA.
ID AC Q8WNR5; AC
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome P450 2D.
OS Macaca fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9542;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Hichiya H., Yamamoto S., Asaoka K., Narimatsu S.;
RT "Molecular cloning and functional analysis of a Japanese monkey CYP2D
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF301911; AAL73443.1; --
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR0385; P450.

DR PROSITE; PS00086; CYTOCHROME P450; 1.
SQ SEQUENCE 497 AA; 55895 MW; C7EECB2ADB74A38 CRC64;

Alignment Scores:

Pred. No.: 2.88e-138 Length: 497
Score: 2138.50 Matches: 417
Percent Similarity: 85.71% Conservative: 9
Best Local Similarity: 83.90% Mismatches: 20
Query Match: 74.10% Indels: 51
DB: 6 Gaps: 1

US-09-820-788A-1 (1-1537) x Q865W1 (1-497)

```
QY 78 ATGGGGCTAGAGCACTGGTCCCTGGCGCTGATAGTGGCACTTCTCTCTCTGCTG 137
Db 1 MetGlyLeuAspAlaLeuValProLeuAlaValThrValAlaIlePheValLeuVal 20
QY 138 GACCTGATGACCGCGCGCAACGCTGGGCTGCACGCTACTCACCAGGCGCCCTGCCACTG 197
Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProGlyProMetProLeu 40
QY 198 CCGGGCTGGGCAACTCTGCTGATGTGCTTCCAGAACACACCATATCTGCTTCGACGAG 257
Db 41 ProPheLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProAsnSerPheAsnGln 60
QY 258 TTGGCGCGCGCTTGGGGAGCGTGTTCAGCTGCGAGCTGGCTGGACCGCGTGGTCTG 317
Db 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpProValValVal 80
QY 318 CTCAATGGCTGGCGCGCGCTGGCGAGCGCTGTGACCCAGCGAGGACACCGCGGAC 377
Db 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
QY 378 CGCGCGCTGTGCCCATCACCAGATCTGGTGTGGTGGTGGCGCGTTCCTCAA----- 428
Db 101 ArgProProValProIleThrGlnMetLeuGlyPheGlyProHisSerGlnGlyValPhe 120
QY 428 ----- 428
Db 121 LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgArgPheSerValSerThrLeuArg 140
QY 428 ----- 428
Db 141 AsnLeuGlyLeuGlyLysLysSerLeuGluGlnTrpValThrGluGluAlaAlaCysLeu 160
QY 429 -----GGACGCGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 464
Db 161 CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys 180
QY 465 GCGGTGACGACGTGATCGCTCCCTCACCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 524
Db 181 AlaValSerAsnValIleAlaSerLeuThrCysArgArgArgPheGluThrAspAsnPro 200
QY 525 CGCTTCTTCAGCTGCTGGACCTAGCTCAGGAGGAGCTGAAGAGGAGGAGTGGGCTTCTG 584
Db 201 CysLeuLeuArgLeuLeuAspLeuThrMetGluGlyLeuLysGluGluSerGlyLeuLeu 220
QY 595 GCGAGGTGCTGAATGCTGTCCCTCTCTCGATATCCAGCGCGCTGGCGGACGAGTGC 644
Db 221 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProGlyLeuAlaGlyLysVal 240
QY 645 CTACGCTTCCAAAGGCTTCTGACCCAGCTGGATGAGCTGCTAACTGAGCAGCAGGATG 704
Db 241 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuThrGluHisArgMet 260
QY 705 ACCTGGGACCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 764
Db 261 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
QY 765 AAGGCCAAGGGGAACCTTGAGAGCGCTTCAATGATGAGAACCTGCGCATAGTGGTGGCT 824
Db 281 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 300
```

```
QY 825 GACCTGTCTCTGCGCGGATGGTGCACCACTCGACCACTGGCGCTGGGCTCTCTGCTC 884
Db 301 AspLeuPheSerAlaGlyMetValThrThrSerIleThrLeuAlaTrpGlyLeuLeuLeu 320
QY 885 ATGATCTCTACATCCGATGTGCAGCGCGTGTCCAAAGAGAGATCGACGATGATAGGG 944
Db 321 MetIleLeuHisProAspValGlnArgValGlnGlnGlnLeuAspValIleGly 340
QY 945 CAGGTGGCGGACAGAGATGGGTGACGAGCTCACATGCCCTACACCACTGCGGTGATT 1004
Db 341 ArgValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaAlaIle 360
QY 1005 CATGAGTGGACGCTTTGGGGACATGTCCTCCCTGGGTGTGACCATATCACATCCCGT 1064
Db 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 380
QY 1065 GACATCGAAGTACAGGCTTCCGATCCCTAAGGAAACGACACTCATCACCACCACTGTCA 1124
Db 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnArgPro 400
QY 1125 TCGTGTCTGAAGATAGGCGCTCTGGAGAGAGCCCTTCGCTTCCACCCCGAAGACTTC 1184
Db 401 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
QY 1185 CTGATGCCAGGCGCACTTTGTCAAGCGGAGGCTTCTCGCTTCTCAGCAGGCGCGC 1244
Db 421 LeuAspAlaGlnGlyArgPheValLysProAspAlaPheLeuProPheSerAlaGlyArg 440
QY 1245 CGTGCATCGCTCGGGAGCGCGCTCGCGCGCATGAGCTTCTCTCTTCTTCCCTCCCTG 1304
Db 441 ArgAlaCysLeuGlyGluProArgAlaArgMetGluLeuPheLeuPheThrCysLeu 460
QY 1305 CTGACGACTTCACTCTCTCGTCCGCTCCACTGGACAGCGCGCGCGCGCGCGCGCTC 1364
Db 461 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 480
QY 1365 TTTCCTTCTCTGCTGAGCCCACTCCCTCATGAGCTTGTGTGTGTGTGTGTGTGTGTGT 1415
Db 481 PheAlaPheLeuValSerProSerProTyrGluLeuCysAlaValProArg 497

RESULT 4
ID Q29454 PRELIMINARY; PRT; 500 AA.
AC Q29454;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome P-450 IID.
GN CYP2D.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=pBVL 180; TISSUE=Liver;
RX MBLINR=9301103; PubMed=1396678;
RA Tsunekawa Y., Matsuo Y., Higuchi R., Ichikawa Y.;
RT "Characterization of the cytochrome P-450IID subfamily in bovine
RT liver. Nucleotide sequences and microheterogeneity.";
RL Eur. J. Biochem. 208:739-746(1992).
CC -1. SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; X68481; CAA48501.1; -.
DR PIR; S37284; S37284.
DR HSSP; P00179; 1DT6.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
```

DR PRINTS; PR00385; P450.
 .DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 500 AA; 55921 MW; 4F62F39050E2BED6 CRC64;

Alignment Scores:
 Pred. No.: 6.09e-113 Length: 500
 Score: 1769.00 Matches: 348
 Percent Similarity: 77.40% Conservative: 39
 Best Local Similarity: 69.60% Mismatches: 59
 Query Match: 61.30% Indels: 54
 DB: 6 Gaps: 2

US-09-820-788a-1 (1-1537) x Q29454 (1-500)

QY	78	ATGGGGCTA-----GAAGCACTGGTGGCCCGTGGATAGTGGCCATCTTCCTG	128
DB	1	MetGlyLeuLeuSerGlyAspThrLeuGlyProLeuAlaValAlaLeuLeuLeuPheLeu	20
QY	129	CTCTGGTGACCTGATGACCGGCCCAACGCTGGGCTGACGCTACTCACCAGGCCCC	188
DB	21	LeuLeuLeuAspLeuMetHisArgArgSerArgTyrAlaProArgTyrProGlyPro	40
QY	189	CTGCCACTCGCGGCTGGCAACCTGCTGCATGCTGCACTTCAGAACACACCATCTGC	248
DB	41	ThrProLeuProValLeuGlyAsnLeuLeuGlnValAspPheGluAspProArgProser	60
QY	249	TTGACCACTGGCGCGCTGGGGACGCTTCAGCTGCAGCTGCAGCTGGCGCTGGACCG	308
DB	61	PheAsnGlnLeuArgArgPheGlyAsnValPheSerLeuGlnValTrpThrPro	80
QY	309	GTGGTCTGCTCAATGGGCTGGCGCGCTGGCGAGCGCTGTGACCCACGCGCAGGAC	368
DB	81	ValValValLeuAsnGlyLeuAlaValArgGluAlaLeuValTyrArgSerGlnAsp	100
QY	369	ACGCGCAGCGCGCTGTGCCATCACCAGATCTCTGGTTTGGCGGCTGCCCA	428
DB	101	ThrAlaAspArgProProAlaValTyrGluHisLeuGlyTyrGlyProArgAlaGlu	120
QY	428	-----	428
DB	121	GlyValLeuAlaArgTyrGlyAspAlaTrpAlaGluGlnArgArgPheSerLeuThr	140
QY	428	-----	428
DB	141	ThrLeuArgAsnPheGlyLeuGlyLysSerLeuGluGlnTrpValThrGluAla	160
QY	429	-----GGACGCCCTTTGGCCCAACGGTCTC	455
DB	161	SerCysSerCysAlaAlaPheAlaAspGlnAlaGlyArgProPheSerProMetAspLeu	180
QY	456	TTGGACAAGCGCTGAGCAACGTGATCGCTCCCTCACCCTGGGGCGCGCTTCGAGTAC	515
DB	181	LeuAsnLysAlaValSerAsnValIleAlaSerLeuThrPheGlyCysArgPheGluTyr	200
QY	516	GACGACCTCGCTTCTCTAGGCTGTGGACCTAGCTCAGGAGGAGTGAAGGAGGAGTGC	575
DB	201	AsnAspProArgIleIleLysLeuLeuAspLeuThrGluAspGlyLeuGluPro	220
QY	576	GGCTTTCTCGGAGGTGAATGCTGTCTCCCGTCTCTCTCCATATCCACGCTGGCT	635
DB	221	AsnLeuValArgLysValValGluAlaValProValLeuLeuSerIleProGlyLeuAla	240
QY	636	GGCAGGTCTAGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGTGTACTGAG	695
DB	241	AlaArgValPheProAlaGlnLysAlaPheMetAlaLeuIleAspGluLeuAlaGlu	260
QY	696	CACAGGATCCTGGACCCAGCCCGGAGCTCAGCTCAGGCTTCTCTGGCA	755
DB	261	GlnLysMetThrArgAspProThrGlnProProArgHisLeuThrAspAlaPheLeuAsp	280
QY	756	GAGATGAGAGGCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACCTGGGCATA	815
			815

DB	281	GluValLysGluAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgLeu	300
QY	816	GTGGTGGCTGACCTGTCTCTCCGGGATGGTGCACACCTGCACACGCTGGCGTGGGC	875
DB	301	ValValAlaAspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTrpAla	320
QY	876	CTCTGCTCATGATCTCATCTCCGATGTGACGCGGTGTCCAAAGAGATCGACGAC	935
DB	321	LeuLeuLeuMetIleLeuHisProAspValGlnArgValGlnGlnGlnIleAspGlu	340
QY	936	GTGATAGGCGAGTGGCGGACGACAGATGGTGGACCGGCTCAGATGCCCTACACCAT	995
DB	341	ValIleGlyGlnValArgArgProGluMetGlyAspGlnAlaLeuMetProPheThrVal	360
QY	996	GCGGTGATTCATGAGTGCAGCGCTTGGGACATCTGCTCCCTGGGTGTGACCCATATG	1055
DB	361	AlaValValHisGluValGlnArgPheAlaAspIleValProLeuGlyLeuProHisMet	380
QY	1056	ACATCCCTGACATCGAAGTACAGGCTTCGCGATCCCTTAAGGGAACGACATCATCAC	1115
DB	381	ThrSerArgAspIleGluValGlnGlyPheHisIleProLysGlyThrThrLeuIleThr	400
QY	1116	RACCTGTCTATCGTCTGAGGATGAGCGCTCTGGAGAACCCCTTCGGCTTCACCC	1175
DB	401	AsnLeuSerSerValLeuLysAspGluThrValTrpGluLysProPheArgPheHisPro	420
QY	1176	GAACACTTCTGGATGCCAGGCCACTTGTGAAGCGGAGGCTTCTGCTCTTCTTCA	1235
DB	421	GluHisPheLeuAspAlaGlnGlyArgPheValLysGlnGluAlaPheIleProPheSer	440
QY	1236	CGACGCCCGCTGATGCTCTGGGAGCCCTTGGCGGATGGAGCTTCTTCTCTTCTTC	1295
DB	441	AlaGlyArgArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhe	460
QY	1296	ACCTCCCTGCTGCAGCACTTCACTTCTCGGTGCCACTTGCAGACGCCCGCGCCAGCC	1355
DB	461	ThrSerLeuLeuGlnHisPheSerPheSerValProAlaGlyGlnProArgProSerGlu	480
QY	1356	CATGTGTCTTGTCTTCTGTCGACCCCATCCCTATGAGCTTGTGCTGCTGCCCGC	1415
DB	481	HisGlyValPheAlaPheLeuValThrProAlaProTyrGlnLeuCysAlaValProArg	500

RESULT 5
 Q9TUJ4
 ID Q9TUJ4 PRELIMINARY; PRT; 500 AA.
 AC Q9TUJ4;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Cytochrome P450 2D/II.
 OS Oryctolagus cuniculus (Rabbit)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=New Zealand White; TISSUE=Liver;
 RX MEDLINE=98391821; PubMed=972858;
 RA Yamamoto Y., Ishizuka M., Takada A., Fujita S.;
 RT "Cloning, tissue distribution, and functional expression of two novel
 rabbit cytochrome P450 isozymes, CYP2D3 and CYP2D4.";
 RL J. Biochem. 124:503-508 (1998).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AB008785; BAA84473.1; -.
 DR PIR; J0258; J0258.
 DR HSSP; P00179; 1DT6.
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01686; EP450ICYP2D.

DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme: Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 55604 MW; 434EC7C86FE305B CRC64;

Alignment Scores:

Pred. No.: 1.04e-111 Length: 500
Score: 1751.00 Matches: 348
Percent Similarity: 75.80% Conservative: 31
Best Local Similarity: 69.60% Mismatches: 67
Query Match: 60.67% Indels: 54
DB: 6 Gaps: 2

US-09-820-788A-1 (1-1537) x Q9TJ4 (1-500)

```
QY 78 ATGGGGCTA-----GAAGCATGTGGTCCCTGGCGGTGATAGTGGCATCTTCTG 128
Db 1 MetGlyLeuSerGlyGluAlaLeuAlaProLeuAlaValAlaValAlaPheLeu 20
QY 129 CTCCTGGTGGACTGATGCACCGCGCAACGCTGGCTGCACGCTACTCCACGAGCC 188
Db 21 LeuLeuValAspLeuMetHisLysArgProArgTrpAlaAlaArgTrpProGlyPro 40
QY 189 CTGCCACTGCCGGCTGGCAACCTGCTGCATGTGGACTTCCAGAACACACCACTACTGC 248
Db 41 ValGlyIleProGlyLeuGlyAsnLeuLeuGlnValAspPheArgGlyIleProAsnCys 60
QY 249 TTCGACAGTTCGGCGCGCTTCGGGACGTGTTCAGCTGCAGCTGCCTGGCTGCAGCGCG 308
Db 61 PheArgGlnLeuArgArgTrpGlyAspValPheSerLeuGlnLeuAlaTrpThrPro 80
QY 309 GTGGTCGTCAATGGCTGCGCGCTGCGCGAGGCGGTGGTACCCACGCGGAGGAC 368
Db 81 ValValValLeuAsnGlyProAlaValIleArgGluAlaLeuValThrTrpGlyLysAsp 100
QY 369 ACCGCGGACCGCGCGCTGTGCCATCACCCAGATCCCTGGGTTCGGCGCGGTTCCTCAA 428
Db 101 ThrAlaAspArgProAlaHisThrLeuGluProLeuGlyPheGlyProHisAlaGln 120
QY 428 ----- 428
Db 121 GlyValValMetAlaArgTrpGlyProAlaTrpArgGluGlnArgPheSerValSer 140
QY 428 ----- 428
Db 141 ThrLeuArgAsnPheGlyLeuGlyLysSerLeuGluGlnTrpValThrGluAla 160
QY 429 -----GGACGCCCTTTTCGCCCCCAACGGTCTC 455
Db 161 ThrCysLeuCysAlaAlaPheAlaAspHisAlaGlyCysProPheSerProSerMetLeu 180
QY 456 TTGGACAAGCGTGAGCAAGTGATCGCTCCCTCAGCTGGCGGCGCGCTTCGAGTAC 515
Db 181 LeuAsnLysAlaValCysAsnValIleAlaSerLeuThrHisGlyCysArgPheGluTrp 200
QY 516 GACGACCTCGCTTCTCAGCTGCTGACCTAGCTCAGGAGGAGTCAAGAGGAGTCTG 575
Db 201 AspAspHisArgLeuThrArgLeuMetAspLeuThrGlnThrIleLeuLysGluSerThr 220
QY 576 GCCTTTCGCGGAGGTGTAATGCTGTCGCCGCTCCCTGCATATCCACGCGCTGGCT 635
Db 221 GlyAsnLeuProGlnValLeuAsnValIleProIleLeuLeuAspGlyIleProGlyLeuVal 240
QY 636 GCACAGTCTACGCTTCCAAAGCGCTTCTGACCCAGCTGGATGAGCTGCTAAGTAC 695
Db 241 AspLysValPheArgGlyGlnLysAlaPheMetAlaLeuLeuAspGluLeuValThrGlu 260
QY 696 CACAGATGACTGGGACCGACCCAGCCCGCCGAGACCTGACTGAGGCTTCTTGCA 755
Db 261 HisArgMetThrArgAspProAlaGlnProProArgAspLeuThrAspAlaPheLeuAsp 280
QY 756 GAGATGGAGAGGCCAAGGGAACCTCGAGAGCAGCTTCAATGATGAGAACCTCGCAT 815
Db 815 ----- 815
```

```
Db 281 GlnValGluLysAlaLysGlyAsnProGluSerSerPheAsnAspAspAsnLeuArgLeu 300
QY 816 GTGTGGCTGACCTGTTCTCTGCGGGATGCTGACCACTGACCACTGCGCTGGCGG 875
Db 301 ValValThrAspLeuPheAlaAlaGlyMetValThrSerIleThrLeuSerTrpAla 320
QY 876 CTCCTGCTCATGATCTTACATCCGATGTCAGCGCGCTGTCCAAACAGGAGATCGAC 935
Db 321 LeuLeuLeuMetIleLeuHisProAspValGlnArgValGlnGlnIleAspGlu 340
QY 936 GTGATAGGCGAGGTGCGGACACAGAGATGGTGCAGAGCTCATGCTCCCTACACACT 995
Db 341 ValIleGlyProAlaArgArgProGluMetGlyAspGlnAlaArgMetProTrpThr 360
QY 996 GCCGTGATTCATGAGTGCAGCGCTTGGGACATCGTCCCTGGGTGTGACCATATG 1055
Db 361 AlaValValHisGluValGlnArgPheAlaAspIleIleProLeuGlyValProHisGln 380
QY 1056 ACATCCCGCTGACATCGAAGTACAGGCTTCGCGCATCCCTAAGGGAACGACACTCATCACC 1115
Db 381 ThrSerArgAspIleGluValGlnGlyPheLeuIleProLysGlyThrValLeuPheThr 400
QY 1116 AACCTGTCATGCTGTGTAAGATGAGCGCTCTGGGAGAAAGCCCTTCGCTTCCACCCC 1175
Db 401 AsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisPro 420
QY 1176 GAACACTCTCTGGATGCCAGGCGCACCTTGTGAGCGGAGGCTTCTGCTGCTTCTCA 1235
Db 421 GlyHisPheLeuAspAlaGlnGlyArgPheValLysGlnAlaPheMetProPheSer 440
QY 1236 GCAGCGCGCTGTCATGCTCGCGGAGCCCTGCGCGCATGGAGCTTCTTCTTCTTC 1295
Db 441 AlaGlyArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPhePhePhe 460
QY 1296 ACCTCCTGCTGACACTTTCAGCTTCTGCTGCTGCTGCTGACAGCCCGCGCGCAC 1355
Db 461 ThrCysLeuLeuGlnArgPheSerValProThrGlyGlnProArgProSerAsp 480
QY 1356 CATGCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1415
Db 481 GlnGlyAlaProAlaThrLeuValThrProAlaProTrpGlnLeuCysAlaValAlaArg 500
```

RESULT 6

Q16804 PRELIMINARY; PRT; 373 AA.
AC Q16804; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome P450db1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89153788; PubMed=2466049;
RA Manns M.P., Johnson E.F., Griffin K.J., Tan E.M., Sullivan K.F.;
RT "Major antigen of liver kidney microsomal autoantibodies in idiopathic
autoimmune hepatitis is cytochrome P450db1.";
RL J. Clin. Invest. 83:1066-1072(1989).
CC -; SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; M24499; AAA36403.1; -;
DR HSSP; P00179; 1DT6.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.

DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
FT NON_TER 1
SQ SEQUENCE 373 AA; 42005 MW; 9FDED67B0BA487A4 CRC64;
Alignment Scores:
Pred. No.: 1.42e-109 Length: 373
Score: 1719.50 Matches: 333
Percent Similarity: 93.82% Conservative: 1
Best Local Similarity: 93.54% Mismatches: 3
Query Match: 59.58% Indels: 19
DB: 4 Gaps: 1
US-09-820-788A-1 (1-1537) x Q16804 (1-373)
QY 405 CTGGTCTTTGGCGCGCTTCC----- 425
Db 18 LeuGlyLeuGlyLysLysSerLeuGluGlnTrpValThrGluAlaAlaCysLeuCys 37
QY 426 -----CAAGACGCCCTTTTCGCCCAACGGTCTCTTGGACAAAGCC 467
Db 38 AlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLysAla 57
QY 468 GTGAGCAAGTGTATCGCTCCCTCACCTGGCGCGCGCTTCGAGTACGACGACCTCGC 527
Db 58 ValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTrpAspAspProArg 77
QY 528 TTCTCTCAGGCTGTGACCTAGCTCAGGAGGAGTCAAGGAGGAGTCTTCGTGCGC 587
Db 78 PheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeuArg 97
QY 588 GAGGTCTGAATGCTGTCCCGCTCTCTGCATATCCAGCGCTGGCTGGCAAGTCTCTA 647
Db 98 GluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysValLeu 117
QY 648 CGCTTCCAAAGCTTCTGACCCAGCTGGATGAGTCTGAAGTCACTGAGCAGCAGGATGACC 707
Db 118 ArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMetThr 137
QY 708 TGGGACCCAGCCAGCCCGCCGAGACCTGACTGAGCGCTTCTGCGCAGAGATGGAGAAG 767
Db 138 TrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGluLys 157
QY 768 GCCAAGGGAACTCTGAGACGAGCTTCAATGATGAGAACCTGCGCATGTGGTGGCTGAC 827
Db 158 AlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAlaAsp 177
QY 828 CTGTTCTCTCCGGGATGTGACCACTCGACACGCTGGCGCTTCTGCTCATG 887
Db 178 LeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeuMet 197
QY 888 ATCTACATCCGATGTGAGCGCGCTGTCCAACAGGAGATCGACGACCTGATAGGCGAG 947
Db 198 IleLeuHisProAspValGlnArgValGlnGlnGluLeuAspValIleGlyGln 217
QY 948 GTGCGCGGACGAGATGGGTGACACGAGCTCATGCTCCCTACACCACTCCGCTGATTCAT 1007
Db 218 ValArgArgProGluMetGlyAspGlnAlaHisMetProTyThrThrAlaValIleHis 237
QY 1008 GAGTTCAGCGCTTTGGGACATGTCGCCCTGGTGTGACCAATATGATCCCGTGAC 1067
Db 238 GluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArgAsp 257
QY 1068 ATCGAAGTACAGGCTTCGCGATCCCTAAGGGAACGACACTCATCACCACTGTCATCG 1127
Db 258 IleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSerSer 277
QY 1128 GTGCTCAAGGATGAGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCCAACACTTCCCTG 1187
Db 278 ValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPheLeu 297
QY 1188 GATGCCACAGGGCACTTTGTGAAGCGGAGGCGCTTCTCCCTGCTTCTACAGCGCGCGCT 1247

Db 298 AspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArgArg 317
QY 1248 GCATGCTCGGGAGAGCCCTGGCCGCGCATGGAGCTTCTCTTACCTCCCTCCCTGCTG 1307
Db 318 AlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeuLeu 337
QY 1308 CAGCACTTCAGCTTCTCGGTGCCCACTGGACAGCCCGCCGCGCACCATGGTGTCTTT 1367
Db 338 GlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyValPhe 357
QY 1368 GCTTTCTCTGCTGACCCCATCCCTATGAGCTTCTGCTGCTGCCCGC 1415
Db 358 AlaPheLeuValThrProSerProTyThrGluLeuCysAlaValProArg 373
RESULT 7
Q9JKY7 PRELIMINARY; PRT; 500 AA.
AC Q9JKY7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Cytochrome P450 CYP2D2.
GN CYP2D2
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20485155; PubMed=11032406;
RT Blume N., Leonard J., Xu Z.J., Watanabe O., Remotti H., Fishman J.;
RA "Characterization of Cyp2d22, a novel cytochrome P450 expressed in
RT mouse mammary cells";
RL Arch. Biochem. Biophys. 381:191-204 (2000).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF221525; AAF34652.1; -.
DR HSSP; P00179; 1D76.
DR MGD; MGI:1929474; Cyp2d22.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 56467 MW; FACB35854CBC3F1F CRC64;
Alignment Scores:
Pred. No.: 5.75e-109 Length: 500
Score: 1711.00 Matches: 337
Percent Similarity: 73.32% Conservative: 34
Best Local Similarity: 66.60% Mismatches: 77
Query Match: 59.29% Indels: 58
DB: 11 Gaps: 2
US-09-820-788A-1 (1-1537) x Q9JKY7 (1-500)
QY 51 AGGACCCATTTGTAGTAGGAGGAGTATGGGCTAGAGCACTGGTCCCTCGCCGTG 110
Db 2 ArgLeuProThrGlyAlaGlu-----LeuTrpProIleAlaIle 14
QY 111 ATAGTGGCATCTTCTGCTGCTGCTGACCTGATGACCGCGCGCAAGCTGGGTGCA 170
Db 15 PheThrValIlePheLeuIleLeuValAsnLeuMetHisTrpArgGlnArgTrpThrAla 34
QY 171 CGCTACTCACAGCGCCCTCTGCACTGCCCGGGCTGGGCAACTCTGCTGATGGACTTC 230
Db 35 HisTyProGlyProMetProTrpProValLeuGlyAsnLeuLeuHisMetAspPhe 54
QY 231 CAGAAACACCATACTGCTTTCACACAGTTCGGCGCCCGCTTCGGGAGCGTGTTCAGCTG 290
Db 55 GlnAsnMetProAlaGlyPheGlnLysLeuArgGlyArgTyArgGlyAspLeuPheSerLeu 74

Qy	291	CAGCTGGCCTGGACCCCGGTGGTCTGTCTCAATGGCTGGCGCGCTGGCGAGCGCTG	350
Db	75	GlnLeuAlaSerGluSerValValValLeuAsnGlyLeuThrAlaLeuArgGluAlaLeu	94
Qy	351	GTGACCCACGGCGAGGACACCGCGACCGCGCGCTGGCCCATCACCCAGATCCTGGT	410
Db	95	ValLysHisSerGluAspThrAlaAspArgProLeuHisPheAsnAspLeuGly	114
Qy	411	TTTGGCGCGCTTCCCAA-----	428
Db	115	PheGlyProArgSerGlnGlyLleValLeuAlaArgTyrGlyProAlaTrpArgGlnGln	134
Qy	428	-----	428
Db	135	ArgArgPheSerValSerThrMetHisHisPheGlyLeuGlyLysLysSerLeuGluGln	154
Qy	429	-----GGAGCGCCC	437
Db	155	TrpValThrGluGluAlaArgCysLeuCysAlaAlaPheAlaAspHisThrGlyHisPro	174
Qy	438	TTTCGCCCAACGGTCTCTTGGACAAACCGGTGAGCAACGTGATCGCTCCCTCACCTGC	497
Db	175	PheSerProAnThrLeuLeuAspLysAlaValCysAsnValIleAlaSerLeuLeuTyr	194
Qy	498	GGGCGCGCTTCGAGTAGACGACGCCCTCGCTTCCTCAGCTGCTGGACCTAGCTCAGAG	557
Db	195	AlaCysArgPheGluTyrAspAspProArgPheIleArgLeuLeuGlyLeuLysGlu	214
Qy	558	GCAGTGAAGGAGTCCGGCTTTCTGCGCGAGGTGCTGAATGCTGCCGCTCCTCCTG	617
Db	215	ThrLeuLysGluGluAlaGlyPheLeuProMetPheLeuAsnValPheProMetLeuLeu	234
Qy	618	CATATCCGAGCGCTGGCTGGCAAGTCTACGTTCCAAAGGCTTTCTTCACCCAGCTG	677
Db	235	ArgIleProGlyLeuValGlyLysValPheProGlyLysArgAlaPheValThrMetLeu	254
Qy	678	GATGAGCTGCTAACTGACACAGGATGACCTGGGACCCAGCCAGCCCCCGAGACCTG	737
Db	255	AspGluLeuLeuAlaGluHisLysThrTrpAspProThrGlnProProArgAspLeu	274
Qy	738	ACTGAGGCTTCCTGGCAGATGGAGAGGCCAAGGGAAACCTTGAGAGCAGCTTCAT	797
Db	275	ThrAspAlaPheLeuAlaGluValGlyLysAlaLysGlyAsnProGluSerSerPheAsn	294
Qy	798	GATCAGAACCTGCGCATAGTGGCTGCTGCTACCTGTTCTCTGCGGGATGGTGACCACTCG	857
Db	295	AspGluAsnLeuArgThrValValGlyAspLeuPheSerAlaGlyMetValThrThrSer	314
Qy	858	ACCACTGGCTGGCGCTCCTGCTCATGATCTCATCTCGGATGTGACGCGCGTCTC	917
Db	315	ThrThrLeuSerTrpAlaLeuMetLeuMetIleLeuTyrProAspValGlnArgVal	334
Qy	918	CAACAGAGATCGACGAGTGTATAGGCGAGTGGCGGACACAGAGATGGGTACAGGCT	977
Db	335	GlnGlnGluIleAspGluValIleGlyGlnValGlnCysProGluMetalAspGlnAla	354
Qy	978	CACATGCCCTACACACTGCCGTGATTATCATGAGTGCAGCGCTTTGGGACATCGTCCC	1037
Db	355	ArgMetProTyrThrAsnAlaValIleHisGluValGlnArgPheAlaAspIleLeuPro	374
Qy	1038	CTGGGTGTGACCCATATGACATCCCGTGCATATCGAAGTACAGGCTTCCCGCATCCCTAAG	1097
Db	375	LeuGlyValProHisLysThrSerArgAspIleGluLeuGlnGlyPheLeuIleProLys	394
Qy	1098	GGAAAGCACTCATCAACAACCTGTATCGGTGCTGAAGGATGAGCGCGTCTGGAGAAG	1157
Db	395	GlyThrThrLeuIleThrAsnLeuSerSerAlaLeuLysAspGluThrValTrpGluLys	414
Qy	1158	CCCTTCGCTTCCACCCCGAACCTTCTTGATGCCAGGCGCACTTGTGAAGCCCGGAG	1217
Db	415	ProLeuCysPheHisProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGlu	434
Qy	1218	GCCTTCCTGCTTCTCAGCAGGCGCGCTGCATGCTCTCGGGGAGCGCTGGCCCGCATG	1277

```
QY 249 TTCGACGAGTTCGGCGCGCTTCGGGACGCTGTTTCAGCTGACCTGCGCTGACGCGC 308
DB 1 PheArgGlnLeuArgCysArgTyrGlyAspValPheSerLeuGlnLeuAlaTrpThrPro 80
QY 309 GTGTCGTCTCAATGGCTGCGCGCTGCGCGAGCGCTGGTGACCCACGGCGAGGAC 368
DB 81 ValValValLeuAsnGlyProAlaAlaMetArgGluAlaLeuValThrTyrGlyGluAsp 100
QY 369 ACCGCGGACCGCGCTGTCGCCATCACCAGATCTCGGGTTTGGCGCGCTGCCAA 428
DB 101 ThrAlaAspArgProTyrSerLeuSerLeuGluHisLeuGlyPheGlyProGlnAlaGln 120
QY 428 ----- 428
DB 121 GlyValIleMetAlaCysTyrGlyHisAlaTrpArgGluGlnArgPheSerValSer 140
QY 428 ----- 428
DB 141 ThrLeuArgAsnPheGlyMetGlyLysLysSerLeuGluHisTrpValThrGluGluAla 160
QY 429 -----GGACGCGCTTTCGCGCCCAACGGTCTC 455
DB 161 AlaCysLeuCysAlaValPheSerGluHisAlaGlyHisProPheSerProLysAlaLeu 180
QY 456 TTGGACAAAGCGCTGAGCAACGTGATCGCTCCCTCACCTCGCGGCGCGCTTCGAGTAC 515
DB 181 LeuAsnLysAlaIleGlyAsnValIleAlaSerLeuThrPheGlyCysArgPheGluTyr 200
QY 516 GACGACCTCGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGACTGAAGGAGGATGTCG 575
DB 201 AspAspHisArgLeuThrArgLeuMetAspLeuIleGluIleMetLeuGluGluSerThr 220
QY 576 GGCCTTTCGCGGAGTCTGAATGCTGCTCCCTGCTCCCTGCTCATATCCAGCGCTGCT 635
DB 221 GlyIleLeuProLeuValLeuAsnValIleProIleLeuLeuArgIleProGlyLeuVal 240
QY 636 GCGAAGTCTACGCTTCCAAAGGCTTCTGACCCAGCTGGATGAGCTGTAACCTGAG 695
DB 241 AspyLysValPheHisGlyGlnLysAlaPheMetAlaLeuLeuAspGluValThrGlu 260
QY 696 CACGAGTACCTGGGACCCAGCCAGCCCGGAGACCTGACTGAGGCGCTTCTTGSCA 755
DB 261 HisArgMetThrArgAspProAlaGlnProProArgAspLeuThrAspAlaPheLeuAsp 280
QY 756 GAGATGAGAGGCCAAGCGGAACCTCAGAGCAGCTTCATGATGACAACTCGGCATA 815
DB 281 GlnValGluLysAlaLysGlyAsnProGluSerSerPheAsnAspAspAsnLeuArgLeu 300
QY 816 GTGTGGCTGACCTGCTCTGCGCGGATGTTGACCACTCGACCACTGCGCTGCGCGGC 875
DB 301 ValValAlaAspLeuPheValAlaGlyMetPheThrThrSerPheThrLeuSerTrpAla 320
QY 876 CTCCTGCTCATGATCTACATCCGATGTCAGCGCTGTCGACGCTGTCACACAGAGATCGAC 935
DB 321 LeuLeuLeuMetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspGlu 340
QY 936 GTGATAGGCGAGTCCGCGGACGAGATGGTGACCGAGGCTCAGTCCCTACCACT 995
DB 341 ValIleGlyProAlaArgProGluMetGlyAspGlnAlaArgMetProTyrThrThr 360
QY 996 GCGGTGATTCAGGTGACGCGCTTTCGGGACATCGTCCCTCGGGTGTGACCCATATG 1055
DB 361 AlaValValHisGluValGlnArgPheAlaAspIleValProLeuGlyValProHisGln 380
QY 1056 ACATCCCTGATCATCGAAGTACAGGCTTCGCGCATCCCTAAGGGGAACGACACTCATCAC 1115
DB 381 ThrLeuArgAspIleGluValGlnGlyPheLeuIleProLysGlyThrMetLeuPheThr 400
QY 1116 AACCTGTCTATCGGTCTCAAGATCAGCGCTCTGGGAGAGCCCTTCGGCTTCACCCC 1175
DB 401 AsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisPro 420
QY 1176 GAACACTTCTGATGCCAGGGCCACTTTGTGAAGCGCGGAGCGCTTCTGCTTCTTCTCA 1235
```

```
DB 421 GlyHisPheLeuAspAlaGlnGlyArgPheValLysGlnGluAlaPheMetProPheSer 440
QY 1236 GCAGGCGCGCTGATGCTCGGGAGCCCTCGGCGAGCTTCGCGCATGGAGCTTCTCTTTC 1295
DB 441 AlaGlyHisArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPhePhe 460
QY 1296 ACCTCCCTGCTGACGACTTCAGCTTCGCTGCGGCTGACAGAGCCCGCGCGCAGCCAC 1355
DB 461 ThrCysLeuLeuGlnArgPheSerPheSerValProAlaGlyGlnProGlnProSerAsp 480
QY 1356 CATGCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1415
DB 481 GlnGlyAlaProAlaThrLeuValThrProAlaProTyrGlnLeuCysAlaValAlaArg 500

RESULT 10
Q8VCX0 PRELIMINARY; PRT; 500 AA.
AC Q8VCX0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to cytochrome P450, 2d9.
GN 1300007K12RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strauberg R.;
RL Submitted (DSC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; BC018344; AAH18344.1; -
DR MGI; MGI:1915694; 1300007K12RIK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; F:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_Cyp2D.
DR Pfam; PF00067; P450; 1
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00867; CYPASE_2; 1.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 56625 MW; 30017F7A1AE4F83E CRC64;

Alignment Scores:
Pred. No.: 1.18e-103 Length: 500
Score: 1633.50 Matches: 315
Percent Similarity: 73.37% Conservative: 46
Best Local Similarity: 64.02% Mismatches: 80
Query Match: 56.60% Indels: 51
DB: 11 Gaps: 1

US-09-820-788A-1 (1-1537) x Q8VCX0 (1-500)
QY 93 CTGTCGCTGCGCGCTGCGATAGTGCCTTCCTGCTCTCTGCTGACCTGATCAGCGG 152
DB 9 LeuTrpProValAlaIlePheThrValIlePheLeuLeuValAspLeuMetHisArg 28
QY 153 CGCCAAAGCTGGGCTGACGCTACTCACAGGCGCCCTGCGCCTGCGCGCTGGGCAAC 212
DB 29 ArgGlnArgTrpThrSerArgTyrProGlyProValProTrpProValLeuGlyAsn 48
QY 213 CTGTCGATGTCCTCCAGAACACCATCTGCTTCGACAGTTCGCGCGCGCTTC 272
DB 49 LeuLeuGlnValAspLeuAspAsnMetProTyrSerLeuTyrLysLeuGlnAsnArgTyr 68
```



```
QY 273 GGGAGCGTTCAGCTGCGAGCTGGCCGAGCCCGGTGCTGCTCAATGGCTGGCG 332
Db 69 GlyAspValPheSerLeuGlnMetAlaTrpLysProValValIleSerGlyLeuLys 88
QY 333 GCGTGGCGGAGCGCTGCTGACCCAGCGAGGACACCGCGAGCCGCGCTGGCC 392
Db 89 AlaValArgGluValLeuValThrCysGlyGluAspThrAlaAspArgProGluMetPro 108
QY 393 ATCACCACGATCGTGGTGGCGCGCTTCCCAAGACGCGCTTTCGCCCAACGCT 452
Db 109 IlePheGlnHisLeuGlyTyrGlyGluLysAlaLysGlyValValPheAlaProTyrGly 128
QY 452 ----- 452
Db 129 ProGluTrpArgGluLeuArgPheSerValSerThrLeuArgAsnLeuGlyLeuGly 148
QY 452 ----- 452
Db 149 LysLysSerLeuGluGlnTrpValThrGluGluAlaGlyHisLeuCysAspAlaPheThr 168
QY 453 -----CTCTGGACAAAGCGGTGAGCAACGTG 479
Db 169 AlaGlnAlaGlySerProLeuAspProTyrThrLeuLeuAsnLysAlaValCysAsnVal 188
QY 480 ATGGCTCCCTCACCTGGCGGCGCGCTTCGAGTACGACGACCCCTGCTTCCTCAGGCTG 539
Db 189 IleAlaSerLeuIleTyrAlaArgPheGluTyrGlyAspProAspPheIleLysMet 208
QY 540 CTGGACCTAGCTCAGGAGGACTGAAGAGGAGTGGGCTTCTCGCGAGGTGCTGAAT 599
Db 209 LeuLysIleLeuLysGluAsnMetGlyGluAsnThrGlyLeuPheProGluValLeuAsn 228
QY 600 GCTGTCCCGCTCTCTCTGCATATCCAGCGCTGGCGCAAGTCTACGCTTCCAAAG 659
Db 229 ThrPheProIleLeuLeuHisIleProGlyLeuAlaAspLysValPheProGlyGlnLys 248
QY 660 GCTTCTCTGACCCAGCTGAGTGTCTAATGACGACGAGTACCTGGACCCAGCC 719
Db 249 ThrPheLeuThrLeuValAsnLysLeuValThrGluHisLysArgThrTrpAspProAsp 268
QY 720 GAGCCCCCGAGACTGACTAGGCTTCCTGGCGAGATGAGAAAGCGGAGAAC 779
Db 269 GlnProProArgAspLeuThrAspAlaPheLeuAlaGluMetGluLysAlaLysGlyAsn 288
QY 780 CTGAGAGCAGCTTCAATCATGACAGACCTGGCATAGTGTGCTGCTCTCTGCC 839
Db 289 ProLysSerPheAsnGluAlaAsnLeuArgLeuValPheAspLeuPheGlyAla 308
QY 840 GGGATGGTGACCACTCGACCACTGGCTGGCGCTCTCTCTCATATCTATCATCCG 899
Db 309 GlyIleValThrSerSerIleThrLeuThrTrpAlaLeuLeuLeuMetIleLeuHisPro 328
QY 900 GATGTGAGCGCGCTGTCACAGAGATCGACGAGTATGAGGAGGAGTGGCGGACCA 959
Db 329 AspValGlnArgArgValGlnGluLeuAspGluValIleGlyGlnValArgCysPro 348
QY 960 GAGATGGGTGACGAGCTCACATGCTACACCTGCGCTGATTCATGAGTGGAGCGC 1019
Db 349 GluMetAlaAspGlnAlaHisMetProTyrThrAsnAlaValIleHisGluValGlnArg 368
QY 1020 TTTGGGACATCGTCCCTGGTGTGACCCATATGACATCCGCTGACATCAAGTACAG 1079
Db 369 PheAlaAspIleValProMetAsnLeuProHisLysThrSerHisAspIleGluValGln 388
QY 1080 GCTTCTCGGATCCCTAAGGAAAGCACTCATCAACCTGTCTCATCGTGTGTAAGGAT 1139
Db 389 GlyPheLeuIleProLysGlyThrThrLeuIleProAsnLeuSerSerThrLeuLysAsp 408
QY 1140 GAGGCGCTGAGAGAGCCCTTCGCTTCCAGCCCGAACACTTCTGTGATCCCGAGGC 1199
Db 409 GluThrValTrpGluLeuProLeuArgPheHisProGluHisPheLeuAspAlaGlnGly 428
QY 1200 CACTTTGTGAAGCGGAGCTTCTCTGCTCTTCTCAGCAGGCGCGCTGATGCTCGG 1259
```

```
Db 429 HisPheValLysProGluAlaPheMetProPheSerAlaGlyArgAlaCysLeuGly 448
QY 1260 GAGCCCTGGCCGCGATGAGGCTCTCTCTTCTTTCACCTCCCTGCTGAGCACTTCAGC 1319
Db 449 GluProLeuAlaArgMetGluLeuPheLeuPheThrCysLeuLeuGlnArgPheSer 468
QY 1320 TTCTCGGTGCGCCACTGGACAGCCCGCGCCAGCCACCATGGTGTCTTCTTCTTCTGGTG 1379
Db 469 PheLeuValProAlaGlyGlnProGlnProSerAspTyrGlyIlePheThrPheLeuVal 488
QY 1380 ACCCATCCCTATGAGCTTGTGCTGTGCCCGC 1415
Db 489 SerProSerProTyrGlnLeuCysAlaPheThrArg 500

RESULT 11
Q8CIM7 PRELIMINARY; PRT; 500 AA.
AC Q8CIM7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RIKEN cDNA 1300006E06 gene.
OS CYP2D6.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC023241; AAH23241.1; -.
DR MGD; MGI:1923529; Cyp2d26.
DR GO; GO:0016712; P:oxidoreductase activity, acting on paired d. .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
SQ SEQUENCE 500 AA; 56975 MW; B3DDCD8DFA3F265 CRC64;

Alignment Scores:
Pred. No.: 4.51e-103 Length: 500
Score: 1625.00 Matches: 317
Percent Similarity: 73.20% Conservative: 49
Best Local Similarity: 63.40% Mismatches: 80
Query Match: 56.31% Indels: 54
DB: Gaps: 2

US-09-820-788A-1 (1-1537) x Q8CIM7 (1-500)
QY 78 ATGGGGCTA-----GAAGCACTGGTGGCCCTGGCGCTCATAGTGGCCATCTTCCTG 128
Db 1 MetGlyLeuLeuValGlyAspLeuTrpAlaValIlePheThrAlaIlePheLeu 20
QY 129 CTCTGTGTGGACCTGATGACCGCGCCCAACGCTGGCGTGTACTCTCACGAGCCCG 188
Db 21 LeuLeuValAspLeuValHisArgGlnArgTrpThrAlaCysTyrProGlyPro 40
QY 189 CTGCCACTGCGCGGCTGGCAACTGCTGCTGATGTGGACTTCCAGAACACACCACTACTGC 248
Db 41 ValProPheProGlyLeuGlyAsnLeuLeuValAspPheGluAsnIleProTyrSer 60
QY 249 TTGACACAGTTGGCGCGCTTCTGGGACGCTTTCAGCTTCCAGCTTGGAGCGCG 308
Db 61 PheTyrLysLeuGlnAsnArgTyrGlyAsnValPheSerLeuGlnMetAlaTrpLysPro 80
QY 309 GTGTGCTGTCTCAATGGGCTGGCGCGCTGGCGAGGCGGTGGTGCACCGCGGAGGAC 368
```

```
Db 81 ValValValValAsnGlyLeuLeuAlaValArgGluLeuLeuValThrTyrGlyGluAsp 100
Qy 369 ACCGCCGACCGCCGCTGCGCCATACACCCAGATCCTCGGTTTGGCCCGCTTCCCAA 428
Db 101 ThrSerAspArgProLeuMetProIleTyrAsnHisIleGlyTyrGlyHisLeuSerLys 120
Qy 429 GGA----- 431
Db 121 GlyValIleLeuAlaProTyrGlyProGluTrpArgGluGlnArgArgPheSerValSer 140
Qy 431 ----- 431
Db 141 ThrLeuArgAspPheGlyLeuGlyLysLysSerLeuGluGlnTrpValThrGluGluAla 160
Qy 432 -----CGCCCTTTCGCCCCCAACGGTCTC 455
Db 161 GlyHisLeuCysAspAlaPheThrLysGluAlaGluHisProPheAsnProSerProLeu 180
Qy 456 TTGGACAAAGCGGTGAGCAAGCTGATCGCTCCCTCCCTGCGGGCGCGCTTCAGTAC 515
Db 191 LeuSerLysAlaValSerAsnValIleAlaSerLeuIleTyrAlaArgArgPheGluTyr 200
Qy 516 GACGACCTCGCTTCTCCTCAGGCTGCTGACCTAGCTCAGGAGGACTGAAGAGGAGTCG 575
Db 201 GluAspProphePheAsnArgMetLeuLysThrLeuLysGluSerLeuGlyGluAspThr 220
Qy 576 GCCTTTCGCGGAGGTCTGATGCTGCTCCCGCTCCTCCATATCCAGCGCTGGCT 635
Db 221 GlyPheValGlyGluValLeuAsnAlaIleProMetLeuLeuHisIleProGlyLeuPro 240
Qy 636 GCGAAGGCTTACGCTTCCAAAGGCTTCTCCTGACCCAGCTGGATGAGCTGCTAACTGAG 695
Db 241 AspLysAlaPheProLysLeuAsnSerPheIleAlaLeuValAsnLysMetLeuIleGlu 260
Qy 696 CACAGGATCACTGGGACCCAGCCAGCCCGCCGAGACCTGACTGAGCGCTTCTTGCA 755
Db 261 HisAspLeuThrTrpAspProAlaGlnProProArgAspLeuThrAspAlaPheLeuAla 280
Qy 756 GAGATGGAGAGGCCAAGGGACCTGAGAGCAGCTTCAATGATGAGAACTGGGATA 815
Db 281 GluValGluLysAlaLysGlyAsnProGluSerSerPheAsnAspLysLeuArgIle 300
Qy 816 GTGGTGGCTGACTGCTTCTCGCGGATGGTACCACCTGACCACTGCGCTGGCGGC 875
Db 301 ValValIleAspLeuPheMetAlaGlyMetValThrThrSerThrLeuSerTrpAla 320
Qy 876 CTCCTGCTCATGCTCATCCGATGTGAGCGCGCTGTCCCAACAGGAGATGACGAC 935
Db 321 LeuLeuLeuMetIleLeuHisProAspValGlnArgArgValHisGlnGluIleAspGlu 340
Qy 936 GTGATAGGCGAGTCCGGGACCCAGAGATGGTGACGAGCTCAGCTCCCTACACCACT 995
Db 341 ValIleGlyHisValArgHisProGluMetAlaAspGlnAlaArgMetProTyrThrAsn 360
Qy 996 GCGGTGATTTCATGAGTGCAGCGCTTTCGGGACATCGTCCCTCGGTGGTGACCCATATG 1055
Db 361 AlaValIleHisGluValGlnArgPheAlaAspIleValProThrAsnLeuProHisMet 380
Qy 1056 ACATCCCGTGACATCAAGTACAGCGCTTCCGCGATCCCTTAAGGGAACGACACTCATACC 1115
Db 381 ThrSerArgAspIleLysPheGlnAspPheIleProLysGlyThrLeuLeuIlePro 400
Qy 1116 AACCTGTGCTGCTGAGATGAGCGGCTTGGGAGAGCCCTTCCGCTTCCACCCC 1175
Db 401 AsnLeuSerSerValLeuLysAspGluThrValTrpGluLysProLeuArgPheTyrPro 420
Qy 1176 GAACACTTCTCGGATGCCAGGCGCACCTTGTGAAGCGGAGCGCTTCTCGCTTTTCA 1235
Db 421 GluHisPheLeuAspAlaGlnGlyHisPheValLysHisGluAlaPheMetProPheSer 440
Qy 1236 CGAGCCCGCGTGCTGCTCGGGAGGCCCTTGGCCCGCATGGAGCTTCTCTCTCTTC 1295
Db 441 AlaGlyArgSerCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePhe 460
```

```
Qy 1296 ACCTCCCTGCTGACGACTTCAGCTTCTCGTGCCTCCACTGACAGCCCGCCGACAC 1355
Db 461 ThrCysLeuLeuGlnArgPheSerPheSerValProAspGlyGlnProArgProSerAsp 480
Qy 1356 CATGCTGCTTGTCTTCTTCTGTCGACCCCATCCCTATGAGCTTGTGCTGCTCCCGC 1415
Db 481 TyrGlyIleTyrThrMetProValThrProGluProTyrGlnLeuCysAlaValAlaArg 500

RESULT 12
Q9DBJ5 PRELIMINARY; PRT; 500 AA.
ID AC Q9DBJ5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE I300006E06Rik protein.
GN CYP2D26 OR I300006E06RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guncich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaishizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AK004915; BAB23666.1; -.
DR HSSP; P00179; 1DT6.
DR MGD; MGI:1923529; Cyp2d26.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01686; EP450CYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 56949 MW; F4C9A03E04C8752D CRC64;
```

```
Alignment Scores:
Pred. No.: 7,24e-103 Length: 500
Score: 1622.00 Matches: 317
Percent Similarity: 73.00% Conservative: 48
Best Local Similarity: 63.40% Mismatches: 81
Query Match: 56.20% Indels: 54
DB: 11 Gaps: 2
```

US-09-820-788A-1 (1-1537) x Q9DBJ5 (1-500)

Qy 78 ATGGGGCTA-----GAAGCACTGGTGGCCCTGCGCGTGATAGTGGCCATCTTCCTG 128

Db	1	MetGlyLeuLeuValGlyAspAspLeuTrpAlaValValIlePheThrAlaIlePheLeu	20
Qy	129	CTCCTGTGGACCTGATGCACCGCGCCAAAGCTGGGCTGCACGCTACTCCACGAGGCC	188
Db	21	LeuLeuValAspLeuValHisArgArgGlnArgTrpThrAlaCysTyrProProGlyPro	40
Qy	189	CTGCACCTGCCCGGGCTGGGCAACTGCTGCATGTGGACTTCCAGAAACACACCACTATGC	248
Db	41	ValProPheProGlyLeuGlyAsnLeuLeuValAspPheGluAsnIleProTyrSer	60
Qy	249	TTGCACACAGTTGCGGCGCGCTTCGGGACGTGTTCACGCTGCAGCTGGCCTGCACGCGC	308
Db	61	PheTyrLysLeuGlnAsnArgTyrGlyAsnValPheSerLeuGlnMetAlaTrpLysPro	80
Qy	309	GTGCTGCTGCTCAATGGGCTCGCGCGCTGGCGCAGGCGTGTGTACCCACGCGAGGAC	368
Db	81	ValValValValAsnGlyLeuLysAlaValArgGluLeuValThrTyrGlyGluAsp	100
Qy	369	ACCGCCAGCCGCGCTGTGCCCATCACCCAGATCCTGGGTTTGGCGCGCTGCCAA	428
Db	101	ThrSerAspArgProLeuMetProIleTyrAsnHisIleGlyTyrGlyHisLysSerLys	120
Qy	429	GGA-----	431
Db	121	GlyValIleLeuAlaProTyrGlyProGluTrpArgGluGlnArgArgPheSerValSer	140
Qy	431	-----	431
Db	141	ThrLeuArgAspPheGlyLeuGlyLysSerLeuGluGlnTrpValThrGluGluAla	160
Qy	432	-----CGCCCTTTCCGCCCAACGGTCTC	455
Db	161	GlyHisLeuCysAspAlaPheThrLysGluAlaGluHisProPheAsnProSerProLeu	180
Qy	456	TTGCACAAAGCGCTGAGCAAGTGAATGCCTCCTCCTCCTCGCGGCGCGCTTCGAGTAC	515
Db	181	LeuSerLysAlaValSerAsnValIleAlaSerLeuIleTyrAlaArgPheGluTyr	200
Qy	516	GACGACCTCGCTTCTCAGGCTCTCGACCTAGCTCAGGAGGACGTGAAGGAGGACTCG	575
Db	201	GluAspProPheAsnArgMetCysLeuLysThrLeuLysGluSerLeuGlyGluAspThr	220
Qy	576	GGCTTTCTGCGCGAGGTGCTGAATGCTGTCTCCCTCCTCCTCGCATATCCAGCGCTGGCT	635
Db	221	GlyPheValGlyGluValLeuAsnAlaIleProMetLeuHisIleProGlyLeuPro	240
Qy	636	GGCAGGTCTACGCTTCCAAAGGCTTCTGTGACCCAGCTGTGATAGCTGTCAACTGAG	695
Db	241	AspLysAlaPheProLysLeuLeuAsnSerPheIleAlaLeuValAsnLysMetLeuIleGlu	260
Qy	696	CACAGGATGACTCGGACCCAGCCAGCCCGCGAGACCTGACTGAGGCTCTCTCGGCA	755
Db	261	HisAspSerThrTrpAspProAlaGlnProProArgAspLeuThrAspAlaPheLeuAla	280
Qy	756	GAGATGAGAGGCCAAGGGGAACCTTGAGACAGCTTCAATGATAGAACCTCGCGATA	815
Db	281	GluValGluLysAlaLysGlyAsnProGluSerSerPheAsnAspLysAsnLeuArgIle	300
Qy	816	GTGGTGGCTGACCTGTTCTCTGCCGGATGGTGAACCTCGACACGCTGGCGCTGGGGC	875
Db	301	ValValIleAspLeuPheMetAlaGlyMetValThrThrSerThrThrLeuSerTrpAla	320
Qy	876	CTCTGCTCATGATCTACATCCGGATGTCAGCGCGCTGCCAACAGGAGATCGACAC	935
Db	321	LeuLeuLeuMetIleLeuHisProAspValGlnArgArgValHisGlnGluIleAspGlu	340
Qy	936	GTGATAGGCGAGGTGCGGCGACACAGATGGGTGACCAAGGCTCATGTCCTACACCACT	995
Db	341	ValIleGlyHisValArgHisProGluMetAlaAspGlnAlaArgMetProTyrThrAsn	360
Qy	996	GCGGTGATCATGAGGTGCAGCGCTTTGGGGACATCGTCCCTCTGGGTGACCATATG	1055

Db	361	AlaValIleHisGluValGlnArgPheAlaAspIleValProThrAsnLeuProHisMet	380
Qy	1056	ACATCCCGTGACATCGAAGTACAGGGCTTCGGCATCCCTAAGGAAACGACACTCATCACC	1115
Db	381	ThrSerArgAspIleLysPheGlnAspPheIleProLysGlyThrLeuIlePro	400
Qy	1116	AACCTGTCACTCGTGCTGAGAGATGAGCGCGTCTGGAGAAAGCCCTTCGCTTCCACCCC	1175
Db	401	AsnLeuSerValLeuLysAspGluThrValTrpGluLysProLeuArgPheTyPro	420
Qy	1176	GAACACTTCCTGGATGCCACGGGCCACTTTGTGAAGCCGAGAGCGCTTCCTGCCTTCTCTCA	1235
Db	421	GluHisPheLeuAspAlaGlnGlyHisPheValLysHisGluAlaPheMetProPheSer	440
Qy	1236	GCAGGCCCGCGTGCATGCTCTCGGGAGCCCTCGCCCGCATGGAGCTCTTCTCTTCTTC	1295
Db	441	AlaGlyArgArgSerCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePhe	460
Qy	1296	ACCTCCCTGCTGCAGCACTTCAGCTTCTCGGTGCCCACTGGACAGCCCCCGCCAGCCAC	1355
Db	461	ThrCysLeuLeuGlnArgPheSerPheSerValProAspGlyGlnProArgProSerAsp	480
Qy	1356	CATGGTGTCTTGTCTTCTCGTGGTGCACCCCATCCCTATGAGCTTTGTGTGTGCCCGCCG	1415
Db	481	TyrGlyIleTyrThrMetProValThrProGluProTyrGlnLeuCysAlaValAlaArg	500
RESULT 13			
ID	Q921V1	PRELIMINARY; PRT; 504 AA.	
AC	Q921V1;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DE	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	Cytochrome P450, 2d9.		
GN	CYP2D9.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RP	{1}		
RN	SEQUENCE FROM N.A.		
RA	Strausberg R.J.		
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
CC	-1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.		
DR	EMBL; BC010593; AAH10593.1; -.		
DR	MCD; MGI-88606; CYP2d9.		
DR	GO; GO:0004497; F:monooxygenase activity; IEA.		
DR	GO; GO:0016712; F:oxygenoreductase activity, acting on paired d. .; IEA.		
DR	GO; GO:0006118; P:electron transport; IEA.		
DR	InterPro; IPR001128; Cytochrome P450.		
DR	InterPro; IPR008069; EP450_CYP2D.		
DR	Pfam; PF00067; P450; 1.		
DR	PRINTS; PR01686; EP450ICYP2D.		
DR	PRINTS; PR0385; P450.		
DR	PROSITE; PS00086; CYTOCHROME P450; 1.		
DR	Heme; Monooxygenase; Oxidoreductase.		
SK	SEQUENCE 504 AA; 56950 MW; 6DC93B3985EFB9A2 CRC64;		

Db	29	ArgGlnArgTrpThrSerArgTrpProProGlyProValProTrpProValLeuGlyAsn	48
Qy	213	CTGCTGCATCTGGACCTTCAGAACACACACACTACTGCTTCGACACAGTTCGCGCCGCTTC	272
Db	49	LeuLeuGlnValAspLeuGlyAsnMetProTrpSerLeuTrpLysLeuGlnAsnArgTrp	68
Qy	273	GGGAGCGTGTTCAGCCCTGCAGCTGGCTGCAGCCCGGTGCTGCTCAATAGGCTGGCG	332
Db	69	GlyAspValPheSerLeuGlnMetAlaTrpLysProMetValValIleAsnGlyLeuLys	88
Qy	333	GCCTGCGCGAGGGCTGTGACCCACCGCGAGGACACCCGCGACCCCGCTGCTGCCCC	392
Db	89	AlaMetLysGluMetLeuLeuThrCysGlyGluAspThrAlaAspArgProProValPro	108
Qy	393	ATCACCCAGATCTGGGTGTTGGCGCGCTTCCCAA-----	428
Db	109	IlePheGluTrpLeuGlyValLysProGlySerGlnGlyValValLeuAlaProTyrGly	128
Qy	428	-----	428
Db	129	ProGluTrpArgGluGlnArgArgPheSerValSerThrLeuArgAsnPheGlyLeuGly	148
Qy	428	-----	428
Db	149	LysLysSerLeuGluAspTrpValThrLysGluAlaAsnHisLeuCysAspAlaPheThr	168
Qy	429	-----CGAGCCCTTTCGCCCAACGGTCTCTTGACAAAGCCGTGACCAACGCTG	479
Db	169	AlaGlnAlaGlnProIleAsnProAsnProMetLeuAsnLysSerThrCysAsnVal	188
Qy	480	ATCGCTTCCCTCACCTCGCGCGCGCTTCGAGTACGACACACCTCGCTTCCCTCAGGCTG	539
Db	189	IleAlaSerLeuIlePheAlaArgArgPheGluTyrGluAspProPheLeuIleArgMet	208
Qy	540	CTGACCTAGCTCAGGAGGACTGAAGGAGAGTTCGGGCTTCTGCGCAGGAGTCTGAAT	599
Db	209	LeuLysValLeuGluGlnSerLeuThrGluValSerGlyLeuIleProGluValLeuAsn	228
Qy	600	GCTGTCCCTCTCTCGTCATATCCAGCGCTGGCTGGCAAGTCTACGCTTCCCAAAG	659
Db	229	AlaPheProIleLeuLeuArgLeuProArgLeuAlaAspLysAlaLeuGlnLys	248
Qy	660	GCTTTCCTGACCCAGCTGGATGAGTCTCTAACTGACACAGGATGACTCGGACCCAGCC	719
Db	249	SerPheIleAlaIleLeuAspAsnLeuLeuThrGluAsnArgThrThrTrpAspProVal	268
Qy	720	CAGCCCCCGAGACTGACTGAGGCTTCTTCTGCGAGAGATGAGAAAGCCAGGGGAAC	779
Db	269	GlnAlaProArgAsnLeuThrAspAlaPheLeuAlaGluIleGluLysAlaLysGlyAsn	288
Qy	780	CCTGAGACGAGCTTCATGATGAGAACCTCGGCATAGTGGTGGCTGACCTGTTCTGCCC	839
Db	289	ProGluSerPheAsnAspGluAsnLeuLeuMetValValArgAspLeuPheGlyAla	308
Qy	840	GGGATGTGACCACTCGACACAGCTGGCTGGGGCTCTCTGCTCATGATCTACATCCG	899
Db	309	GlyMetLeuThrThrSerThrLeuSerTrpAlaLeuMetLeuMetIleLeuHisPro	328
Qy	900	GATGTGAGCGCGGTGTCCAAACAGGAGATCGACGACGTGATAGGGCAGGTGCGCGCACCA	959
Db	329	AspValGlnArgArgValGlnGlnGluIleAspGluValIleGlyGlnValArgHisPro	348
Qy	960	GAGATGGGTGACCAAGGCTCATGCCCTACACCACTCGCTGATTCATGAGGTGCAGCGC	1019
Db	349	GluMetAlaAspGlnAlaHisMetProTrpThrAsnAlaValIleHisGluValGlnArg	368
Qy	1020	TTTGGGACATCGTCCCTCGGGGTGCACCATATGATCATCCCGTGCATCGCAAGTACAG	1079
Db	369	PheGlyAspIleValProValAsnLeuProArgIleThrSerHisAspIleGluValGln	388
Qy	1080	GGCTTCGCGCATCCCTAAGGGAACGACACTCATCAACCACTGTCTCATCGGTCTGAAGGAT	1139

Db 389 AspPheLeuIleProIysGlyThyIleLeuLeuProAsnMetSerSerMetLeuIysAsp 408
 Qy 1140 GAGCGCTCTGGGAGAGCCCTTCGCTTCCACCCCGAAACACTTCTCTGGATGCCAGGCG 1199
 Db 409 GluSerValTrpGluIysProLeuArgPheHisProGluHisPheLeuAspAlaGlnGly 428
 Qy 1200 CACTTTGTGAAGCGGAGGCGCTTCCTGCCTTTTCTCAGCAGGCGCGCTGCATCGCTCGG 1259
 Db 429 HisPheValIysProGluAlaPheMetProPheSerAlaGlyArgArgSerCysLeuGly 448
 Qy 1260 GAGCGCTCGCCCGCATGGAGCTTCTCTCTTCTCACCCTCCCTGCTGCAGCAGCTTTCAGC 1319
 Db 449 GluAlaLeuAlaArgMetGluLeuPhePhePhePhePhePhePhePhePhePhePhePhe 468
 Qy 1320 TTCTCGGTGCCACTGGACAGCGCCGCGCAGCCACATGCTGCTTTGCTTCCTTCCTGGTG 1379
 Db 469 PheSerValProAspGlyGlnProGlnProSerAsnSerGlyValGlyGlyIleLeuVal 488
 Qy 1380 ACCCATCCCGCTATGAGCTTCTGCTGTGTCGCCCGCTAGTAATGG 1424
 Db 489 AlaProSerProTyrGlnLeuCysAlaValValArgAspGlnGly 503
 RESULT 14
 Q64529 PRELIMINARY; PRT; 504 AA.
 AC Q64529
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-JUN-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE P45016a-msl.
 OS Mus spretus (Western wild mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=95349581; PubMed=7623810;
 RA Sueyoshi T., Kobayashi R., Nishio K., Aida K., Moore R., Wada T.,
 RA Handa H., Negishi M.;
 RT "A nuclear factor (NF2d9) that binds to the male-specific P450 (Cyp
 2d-9) gene in mouse liver.";
 RT Mol. Cell. Biol. 15:4158-4166(1995).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR ENBL; U20087; AAC52245.1; -;
 DR PTR; I49427; I49427.
 DR HSP; P00179; I0T6.
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome P450.
 DR InterPro; IPR008069; EP450_CYP2D.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01686; EP450ICYP2D.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 504 AA; 57031 MW; 1D9D1363D8C0C511 CRC64;

Alignment Scores:
 Pred. No.: 3 7e-100 Length: 504
 Score: 1582.50 Matches: 310
 Percent Similarity: 72.32% Conservative: 48
 Best Local Similarity: 62.63% Mismatches: 86
 Query Match: 54.83% Indels: 51
 Gaps: 11

Db . 9 LeuTrpProValAlaIlePheThrValIlePheIleLeuLeuValValLeuThrHisGln 28


```
QY 222 GTGAGCTCCAGAACACACCATACTGCTTCGACAGTTGCGCGCGCTTCGGGACGTG 281
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 52 VallaspLeuAsnAsnMetProTyrSerLeuTyrLysLeuGlnAsnArgTyrGlyAspVal 71
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 282 TTCAGCCTGCAGCTGCGCTGACGCGCGGTGTGCTGCTCAATGGGCTGCGCGCGGTGCGC 341
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72 PheSerLeuGlnMetAlaTyrLysProMetValValIleAsnArgMetLysAlaMetLys 91
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 342 GAGCGCTGTGTACCCACCGGAGAGACACCGCGAGACCGCGCGCTGTGCTCCCATCACCCAG 401
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92 GluValLeuLeuThrCysGlyGluAspThrAlaAspArgProProValProIlePheGlu 111
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 402 ATCTCTGGCTTTGGSCCGCGTTCCCAAGCAGCGCCCTTT----- 440
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112 HisLeuGlyPheLysProArgSerGlnGlyMetIlePheAlaProTyrGlyProGluTrp 131
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 440 ----- 440
Db 132 ArgGluGlnArgArgPheSerLeuSerSerLeuArgAsnPheGlyLeuGlyArgLysSer 151
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 440 ----- 440
Db 152 LeuGluGluTrpValIleLysGluAlaGlyHisLeuCysAspAlaPheThrThrGlnAla 171
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 441 -----CGCCCAACCGTCTCTTGACAAAGCGGTGAGCAACGTGATCGCTCC 488
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 GlyGlnTyrIleAsnProAsnThrMetLeuLysLysAlaThrCysAsnValIleAlaSer 191
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 489 CTCACCTCGCGGCGCGCTTCGAGTACGACGACCTCGCTTCCTCAGCTGCTGACCTA 548
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 192 LeuIlePheAlaArgArgPheGluTyrGluAspProTyrLeuIleArgMetLeuLysVal 211
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 549 GCTCAGGAGGACTGAAGAGAGTTCGGGCTTTCTGCGCGAGGTGCTGAATGCTGTCGCC 608
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 LeuGluAspSerLeuThrGluLeuSerGlyLeuIleProGluValIleAsnThrPhePro 231
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 609 GTCTCTCTGCATATCCACGCGTGTGCGCAAGTCTCTACGCTTCCAAAGGCTTTCTGTG 668
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 IleLeuLeuHisIleProArgLeuAlaAspLysPheLeuGlnSerGlnLysSerPheIle 251
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 669 ACCCAGCTGGATGAGCTGTAACCTGAGCACAGGATGACCTGGGACCCACGCGCGCGCC 728
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 252 AlaIleValAspAsnLeuThrGluAsnArgThrThrTrpAspProAlaGlnSerPro 271
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 729 CGAGACCTGACTAGGCGCTTCTGCGCAGATGAGAGAGCGCAAGGGGAACCTGAGAGC 788
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 ArgAsnLeuThrAspAlaPheLeuAlaGluIleGluLysAlaLysGlyAsnProGluSer 291
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 789 AGCTTCAATGATGAGAACTGCGCATAGTGTGCTGACCTGTTCTCTGCGCGGATGGT 848
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 SerPheAsnAspGluAsnLeuArgMetValIleAspLeuPheThrAlaGlyIleLeu 311
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 849 ACCACCTCGACACGCTGGCGCTTCCTGCTCATGCTCATCATCCGATGTCGAG 908
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 312 ThrThrSerThrThrLeuSerTrpAlaLeuLeuLeuMetIleLeuHisProAspValGln 331
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 909 CGCGGTGTCACACAGGAGATGACGACGTGATAGGCGAGGTGCGCGCACGAGATGGGT 968
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 332 ArgArgValGlnGlnGluIleAspGluValIleGlyGlnValArgHisProGluMetAla 351
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 969 GACCAAGGCTACATGCCCTACACCACTGCCGTGATTATGAGGTGCAGCGCTTTGGGGAC 1028
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 AspGlnAlaHisMetProTyrThrAsnAlaValIleHisGluValGlnArgPheGlyAsp 371
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1029 ATCGTCCCGCTGGGTGACCCATATGATCCCGTGACATCGAAGTACAGGCTTCCGC 1088
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 372 IleValProLeuHisLeuProArgIleThrSerArgAspIleGluValGlnAspPheLeu 391
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1089 ATCCCTAAGGAAACGACACTCATCAACCTGCTCATCGGTGCTGAAGGATGAGGCCGTC 1148
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 392 IleProLysGlyThrIleLeuProAsnMetSerSerValHisMetAspThrVal 411
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1149 TGGGAGAAGCCCTTCGCTTCCACCCCGAACACTTCTCTGGATGCCCGCGGCACCTTTGTG 1208
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: February 25, 2004, 02:56:04
Job time : 94.8984 secs

```
Db 412 TrpGluLysProLeuArgPheHisProGluHisPheLeuAspAlaGlnGlyHisPheVal 431
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1209 AAGCCGAGGCGCTTCTGCGCTTCTCAGCAGGCGCGCTGTCATGCTCGGGAGCCCTG 1268
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 432 LysHisGluAlaPheIleThrPheSerAlaGlyArgSerCysLeuGlyGluProLeu 451
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1269 GCCCGCATGGAGCTTCTTCTTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 1328
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 AlaArgMetGluLeuPheLeuPhePheThrCysLeuLeuGlnArgPheSerPheSerVal 471
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1329 CCCACTGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1388
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 472 ProAspGlyGlnProGlnProSerAspHisArgValPheSerIleMetValAlaProSer 491
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1389 CCCTATGAGCTTTGTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1424
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 492 ProTyrGlnLeuCysAlaValIleArgGluGlnGly 503
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 25, 2004, 02:12:03 ; Search time 64.6542 Seconds
(without alignments)
13433.798 Million cell updates/sec

Title: US-09-820-788A-1

Perfect score: 2886

Sequence: 1 cctgctgtgctctgtgtgcc.....aaaaaaaaaaaaaaaaaaaaa 1537

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cn2_1/USPTO spool_p/US09820788/runat_24022004_141425_20017/app_query.fasta_1.12174
-DB=A_Geneseq_2Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09820788 @CGN 1.1.859 @runat_24022004_141425_20017 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_29Jan04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2330	80.7	446	AAO26404	Human dru
2	2294.5	79.5	502	ABU09595	Human cyt
3	2290.5	79.4	497	ADb25834	Human CYP
4	2290.5	79.4	502	ABU09594	Human cyt
5	2289.5	79.3	497	AAR72376	Human aux
6	2289.5	79.3	497	AAR93183	Human cyt
7	2288.5	79.3	497	ADb25832	Human CYP
8	2286.5	79.2	497	ABU09598	Human cyt
9	2286.5	79.2	497	ADb25831	Human CYP
10	2285.5	79.2	497	AAR72377	Human aux

11	2285.5	79.2	497	2	AAR93184	Aar93184 Human cyt
12	2285.5	79.2	497	2	AAR81462	Aar81462 Human der
13	2285.5	79.2	497	5	ABU09563	Abu09563 Human CYP
14	2285.5	79.2	497	6	AAO26405	Aao26405 Human dru
15	2285.5	79.2	497	6	ADb25833	Adb25833 Human CYP
16	2282.5	79.1	497	6	ABU09593	Abu09593 Human cyt
17	2282.5	79.1	497	6	ABR82026	AbR82026 Human cyt
18	2282.5	79.1	497	7	ADb60558	AdB60558 Human Pro
19	2281.5	79.1	497	2	AAR72375	Aar72375 Human aux
20	2281.5	79.1	497	2	AAR93182	Aar93182 Human cyt
21	2279.5	78.0	497	2	AAW44869	Aaw44869 Cytochrom
22	2277.5	78.9	497	2	AAR72378	Aar72378 Human aux
23	2277.5	78.9	497	2	AAR93185	Aar93185 Human cyt
24	2093.5	72.5	497	4	AAE05171	Aae05171 Human dru
25	1719	59.6	500	7	ADb60556	AdB60556 Rat Prote
26	1619	56.1	504	6	AAO22644	Aao22644 Protein o
27	1114	38.6	436	4	AAU68579	Aau68579 Human nov
28	1011	35.0	195	2	AAW44870	Aaw44870 Cytochrom
29	944	32.7	184	4	AAU74361	Aau74361 Human col
30	857.5	29.7	497	5	AAU91321	Aau91321 Killifish
31	837.5	29.0	502	7	ADD46231	Add46231 Human Pro
32	828.5	28.7	502	7	ADD46229	Add46229 Rat Prote
33	821.5	28.5	497	5	AAU91322	Aau91322 Killifish
34	801	27.8	494	7	ADb63662	AdB63662 Rat Prote
35	797.5	27.6	490	2	AAR72370	Aar72370 Human aux
36	797.5	27.6	490	2	AAR89861	Aar89861 Cytochrom
37	797.5	27.6	490	2	AAR93177	Aar93177 Human cyt
38	797.5	27.6	490	2	AAR81463	Aar81463 Human der
39	797.5	27.6	490	2	AAW64072	Aaw64072 Human cyt
40	797.5	27.6	490	6	ABU96553	Abu96553 Human cyt
41	796.5	27.6	490	6	ABU09971	Abu09971 Human cyt
42	796.5	27.6	508	3	AAU53365	Aau53365 Human col
43	795.5	27.6	490	2	AAR72372	Aar72372 Human aux
44	795.5	27.6	490	2	AAR93179	Aar93179 Human cyt
45	794.5	27.5	490	2	AAR72371	Aar72371 Human aux

ALIGNMENTS

RESULT 1

AAO26404
ID AAO26404 standard; protein; 446 AA.

AC AAO26404;
AC AAO26404;
DT 30-JAN-2003 (first entry)
DE Human drug-metabolising enzyme protein.
KW Cytostatic; antiparkinsonian; gene therapy; cancer; Parkinson's disease;
KW human drug-metabolising protein; enzyme.
XX Homo sapiens.
XX
XX WO2002792333-A1.
XX
PD 10-OCT-2002.
XX
PF 01-APR-2002; 2002WO-US009738.
XX
PR 30-MAR-2001; 2001US-00820788.
XX
PA (PEKE) PE CORP NY.
PA (DFRA/) DI FRANCESCO V.
XX (BRAS/) BEASLEY E M.
XX Shao W, Yan C;
XX WPI; 2003-040649/03.
XX N-PSDB; AAL53565, AAL53566.
XX

New human drug-metabolizing proteins and nucleic acids related to the Cytochrome P450 IID drug-metabolizing enzyme subfamily, useful for

PT treating a condition mediated by a human enzyme protein e.g., cancer.

PS Claim 1; Fig 2A; 72pp; English.

XX The invention relates to a novel isolated polypeptide comprising a 446-
 CC amino acid sequence or its allelic variant, orthologue or fragment. The
 CC allelic variant or orthologue is encoded by a nucleic acid that
 CC hybridises under stringent conditions to the opposite strand of the
 CC nucleic acid comprising a sequence having 1537 or 10278 bp. The fragment
 CC comprises at least 10 contiguous amino acids of the 446-amino acid
 CC sequence. The polypeptide is useful for preparing a pharmaceutical
 CC composition for treating a disease or condition mediated by a human
 CC enzyme protein, e.g. cancer or Parkinson's disease. This sequence
 CC represents the human drug-metabolising protein of the invention

XX Sequence 446 AA;

Alignment Scores:

Pred. No.: 1,946-167 Length: 446
 Score: 2330.00 Matches: 446
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.73% Indels: 0
 DB: 6 Gaps: 0

US-09-820-788A-1 (1-1537) x AAO26404 (1-446)

QY 78 ATGGGGCTAGAGCACTGGTCCCTGGCGGTAGTAGTGCGCATTTCTGCTCTCTGTG 137
 Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
 QY 138 GACCTGATGACCGCGCGCAACGCTGGCTGACCTACTACAGGCGCCCTGCCACTG 197
 Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40
 QY 198 CCGCGCTGGGCAACCTGCTGATGTGGACTTCCAGAACACACCATCTGCTTCACACAG 257
 Db 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
 QY 258 TTGCGCGCGCTTCGGGACCTGTTCAGCTGCGAGCTGGCTGACCGCGGTGTCTGTG 317
 Db 61 LeuArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpProValValVal 80
 QY 318 CTCATGGCTGGCGCGCTGCGGAGCGCTGGTGGCCCGGACCCAGGAGGACCGCGAC 377
 Db 81 LeuAsnGlyLeuAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
 QY 378 CGCCCGCTGTGCCATCACCAGATCCTGGGTTTGGCGCGCTTCCAGAGCGCCCC 437
 Db 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyArgPro 120
 QY 438 TTTCCGCCCAACGGTCTCTTGACAAAGCGGTGAGCAACGTGATCGCTTCCTCCTCCTGC 497
 Db 121 PheArgProAsnGlyLeuLeuAspLysAlaValSerAsnValIleAlaSerLeuThrCys 140
 QY 498 GGGCGCGCTTCAGTACGACACCTCGCTTCCTCAGCTGCTGGACCTAGCTCAGGAG 557
 Db 141 GlyArgPheGlyTyrAspAspProArgPheLeuArgLeuLeuAspLeuAlaGlnGlu 160
 QY 558 GGACTGAAGGAGGAGTGGGCTTTCTGCGGAGGTGGTGAATGCTGTCTCCCTCTCTCTG 617
 Db 161 GlyLeuLysGluLysSerGlyPheLeuArgGluValLeuAsnAlaValProValLeuLeu 180
 QY 618 CATATCCCAACGCTGGCTGGCAAGTCTTACGCTTCCAAAGGCTTTCTTGACCCAGCTG 677
 Db 181 HisIleProAlaLeuAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeu 200
 QY 678 GATGAGCTGCTAACTGAGCACAGGATGACCTGGGACCCAGCCAGCCCGCCCGGACCTG 737
 Db 201 AspGluLeuLeuThrGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeu 220
 QY 738 ACTGAGGCGCTTCTGCGAGATGAGAGGCAAGGGGAACCCCTGAGAGCAGCTTCAAT 797

Db 221 ThrGluAlaPheLeuAlaGluMetGluLysAlaLysGlyAsnProGluSerSerPheAsn 240
 QY 798 GATGAGAACCTGCGCATAGTGGTGGCTGACCTGTTCTCTGCGGATGGTGACACCTCG 857
 Db 241 AspGluAsnLeuArgIleValAlaAspLeuPheSerAlaGlyMetValThrSer 260
 QY 858 ACCACGCTGGCTGGGCGCTCTCTCATGATCTTACATCCGATGTGCAGCGCGGTGTC 917
 Db 261 ThrThrLeuAlaTrpGlyLeuLeuMetIleLeuHisProAspValGlnArgArgVal 280
 QY 918 CAACAGAGATCGACGACGTCATAGGGCAGGTGGCGGACAGAGATGGGTGACAGGCT 977
 Db 281 GlnGlnGluIleAspAspValIleGlyGlnValArgProGluMetGlyAspGlnAla 300
 QY 978 CACATGCCCTACACCACTGCGCTGATTCATGAGTGGCAGCGCTTTGGGGACATCGTCCC 1037
 Db 301 HisMetProTyrThrThrAlaValIleHisGluValGlnArgPheGlyAspIleValPro 320
 QY 1038 CTGGGTGTGACCCATATGACATCCCGTACATCGAAGTACAGGGCTTCCGTCATCCCTAAG 1097
 Db 321 LeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPheArgIleProLys 340
 QY 1098 GGAACGACATCATCAACACCTGTCATCGGTGTGTAAGATGAGCGCTCTGGGAGAAG 1157
 Db 341 GlyThrThrLeuIleThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLys 360
 QY 1158 CCCTTCGCTTCCACCCCGAACACATCTCTGATGCCAGGCGCACTTTGTGAACCGGAG 1217
 Db 361 ProPheArgPheHisProGluHisPheLeuAspAlaGlnGlyHisPheValIleProGlu 380
 QY 1218 GCCTTCCTGCTTCTCAGCAGGCGCGTGCATGCTCGGGAGCCCTCGCGCGCATG 1277
 Db 381 AlaPheLeuProPheSerAlaGlyArgAlaCysLeuGlyGluProLeuAlaArgMet 400
 QY 1278 GAGCTTCTCTCTTTCACCTCCCTGCTGACACTTCAGCTTCTCGGTGCCCACTGGA 1337
 Db 401 GluLeuPheLeuPhePheThrSerLeuLeuGlnHisPheSerPheSerValProThrGly 420
 QY 1338 CAGCCCGCGCCAGCCACCATGGTGTCTTCTGTTCTGTTGTTGACCCCATCCCTATCAG 1397
 Db 421 GlnProArgProSerHisHisGlyValPheAlaPheLeuValThrProSerProTyrGlu 440
 QY 1398 TTTTGTGTGTGCGCCCGC 1415
 Db 441 LeuCysAlaValProArg 446
 RESULT 2
 ABU09595
 ID ABU09595 standard; protein; 502 AA.
 XX AC ABU09595;
 XX 16-JUL-2003 (first entry)
 XX Human cytochrome p450 gene CYP2D6, variant G5799C/C5816AT, protein.
 XX Human; cytochrome P450; CYP2D6; chromosome 22; SNP;
 KW single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
 KW psychiatric disorder; drug sensitivity.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH Misc-difference 486
 FT /note= "Wild-type Ser substituted by Thr, as the result
 FT of a single nucleotide polymorphism"
 FT Misc-difference 492..502
 FT /note= "These amino acids replace the wild-type sequence
 FT (Leu-Cys-Ala-Val-Pro-Arg) as the result of a single-
 FT nucleotide polymorphism which causes a frameshift."
 XX PN EP1281755-A2.

PD 05-FEB-2003.
XX
PF 16-JUL-2002; 2002BP-00254972.
XX
PR 31-JUL-2001; 2001US-0309111P.
XX
PA (PFIZ) PFIZER PROD INC.
XX
PI Milos PM, Webb SM;
XX
DR WPI; 2003-373769/36.
DR N-PSDB; ACA61305.
XX
XX
PT New cytochrome P450 2D6 gene variants and polypeptides, useful for
PT determining if a subject has or is at risk of developing a drug
PT sensitivity condition or disorder that is associated with an aberrant
PT CYP2D6 activity.
XX
XX
PS Claim 4; Fig 8; 88pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising a cytochrome
CC P450 2D6 gene variant, e.g. G5799C or C5816AT (referring to the genomic
CC sequence or the same variant nucleotide in the corresponding cDNA
CC sequences). Also included are probes, primers (allele specific
CC oligonucleotides) and arrays used to detect and/or amplify the CYP2D6
CC gene polymorphic regions, the variant polypeptides, antibodies which are
CC capable of distinguishing between the variant and wild-type polypeptides,
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
CC determining whether an individual is susceptible to being a poor
CC metaboliser of drugs. The DNA probe is useful for hybridising to a
CC variant form of the CYP2D6 gene. The primer is useful for amplifying the
CC C5816TA allelic variant. The allele specific nucleotide is useful for the
CC detection of the C5816TA allelic variant. The methods are useful for
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and
CC determining if an individual is susceptible to being a poor metaboliser
CC of drugs. The nucleic acids are useful as probes or primers for
CC determining whether a subject has a genetic deficiency for metabolising
CC drugs that are substrates of P450 CYP2D6. The methods are useful for
CC determining if a subject has or is at risk of developing a drug
CC sensitivity condition or disorder that is associated with an aberrant
CC CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
CC aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
CC appropriate drugs or determining the course of treatment to administer to
CC a subject to treat cardiovascular or psychiatric disorders, or for
CC treating a subject with a drug sensitivity or disorder associated with a
CC specific allelic variant of a polymorphic region of the CYP2D6 gene. The
CC antibodies are useful for monitoring CYP2D6 protein levels in an
CC individual for determining whether a subject has a disease or conditions
CC associated with an aberrant CYP2D6 protein level. The gene is located on
CC human chromosome 22. The present sequence is the G5799C/C5816AT double
CC variant CYP2D6 protein
XX
SQ Sequence 502 AA;

Alignment Scores:			
Pred. No.:	9,67e-165	Length:	502
Score:	2294.50	Matches:	448
Percent Similarity:	89.44%	Conservative:	1
Best Local Similarity:	89.24%	Mismatches:	2
Query Match:	79.50%	Indels:	52
DB:	6	Gaps:	1

US-09-820-788A-1 (1-1537) x ABU09595 (1-502)

QY 78 ATGGGGCTAGAGCACTGGTCCCTCGCGTAGTGGCCATCTTCTGCTCTCGTG 137
DB 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20

QY 138 GACCTGATGACCGCGCCACAGCTGGGCTGACCGCTACTACAGGCGCCCTGCACTG 197
DB 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProGlyProLeuProLeu 40

QY 198 CCCGGCTGGGCAACCTGCTGATGTGACATTCCAGAAACACACCATATCTGCTTCACCCAG 257
DB 41 ProGlyLeuGlyAsnLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60

QY 258 TTGGGGCCCGCTTCGGGGAGCTGTTTCAGCTCGAGCTGGCTGGAGCCCGGTGGTCTG 317
DB 61 LeuArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValVal 80

QY 318 CTAATGGCTGGGGCCGCTGGCGAGCGCTGGTACCCACGGCGAGACACCCCGAC 377
DB 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaasp 100

QY 378 CGCCCGCTGTGCCCATCACCCAGATCTGGGTTTGGGCGCGTTCCAA 428
DB 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyValPhe 120

QY 428 ----- 428

DB 121 LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgPheSerValSerThrLeuArg 140

QY 428 ----- 428

DB 141 AsnLeuGlyLeuGlyLysLysSerLeuGluGlnTrpValThrGluGluAlaAlaCysLeu 160

QY 429 -----GGACGGCCCTTTCCGCCCAACGGGTCTCTTGGACAAA 464
DB 161 CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys 180

QY 465 GCGGTGAGCAACGATCGCTCCCTACCTCGCGGGCGCGCTTCGACTACGACACCT 524
DB 181 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGlyLeuTyrAspAspPro 200

QY 525 CGCTTCTTCAGCTGCTGGACCTAGCTCAGGAGGAGCTGAAGGAGGAGCTCGGGCTTCTG 584
DB 201 ArgPheLeuArgLeuLeuAspLeuAlaGlnGlyLeuLysGluLysSerGlyPheLeu 220

QY 585 GCGAGGTGCTGAATGTGTCTCCCTGCTGATATCCAGCGCTGGCTGGCAAGTTC 644
DB 221 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 240

QY 645 CTAGCTTCCAAAAGGCTTCTGACCCAGCTGATGAGCTGCTAACTGAGCAGCAGATG 704
DB 241 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 260

QY 705 ACCTGGGACCCAGCCAGCCCGAGCTGAGGCTTCTGAGGCTTCTCGCAGAGATGAG 764
DB 261 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280

QY 765 AAGCCCAAGGGGAACCTCGAGAGCAGCTTCAATGATGAGAACCTCGCATAGTGGTCT 824
DB 281 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 300

QY 825 GACCTGTCTCTGCGGGATGGTGACCACTCGACCGCTGGCGCTGGGGCTCTCTGCTC 884
DB 301 AspLeuPheSerAlaGlyMetValThrSerThrLeuAlaTrpGlyLeuLeuLeu 320

QY 885 ATGATCTTACATCCGATGTCAGCGCTGTCCAAACAGGAGATCGACGCTGATAGG 944
DB 321 MetIleLeuHisProAspValGlnArgValGlnGlnGlnIleAspAspValIleGly 340

QY 945 CAGGTGGCGGCACAGAGATGGGTGACCGAGCTCACATGCCCTTACACCACTGCGCTGATT 1004
DB 341 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 360

QY 1005 CATGAGGTGACGCTTTGGGGACATGTCCTCCCTGGGTGGTGGCCATATGACATCCCGT 1064
DB 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyMetThrHisMetThrSerArg 380

QY 1065 GACATCAAGTACAGGCTTCCCGATCCCTAAGGGAGCAGCACTCATCACCAACTGTCA 1124
DB 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400

Db 341 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 360
QY 1005 CATGAGGTGAGCGCTTTGGGAGCATCGTCCCTGGGTGTGACCCATATGACATCCCGT 1064
Db 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 380
QY 1065 GACATCGAAGTACAGCGCTTCGCATCCCTTACGGGAACGACACTCATCAACCACTGTCA 1124
Db 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400
QY 1125 TCGGTCTCTGAAGATGAGCGCTCTCGGAGAGCCCTTCGGTTCACCCCGAACACTTC 1184
Db 401 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
QY 1185 CTGATGCCAGGCGCACTTTGTGAAGCGGAGCGCTTCCTGCTTTCTCAGAGCGCGC 1244
Db 421 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 440
QY 1245 CGTGCATGCTCGGAGCGCTCGCGCATGAGCTCTCTCTCTTCACTCCCTCGT 1304
Db 441 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 460
QY 1305 CTGCAGCACTTACGCTTCGGTCCCACTTGGACAGCCCGCGCCAGCCACCATGGTGC 1364
Db 461 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 480
QY 1365 TTTCGTTCTCTGTTGACCCATCCCTATGAGCTTTGTGCTGTGCGCCGC 1415
Db 481 PheAlaPheLeuValSerProSerProTyrGluLeuCysAlaValProArg 497
RESULT 4
ABU09594
ID ABU09594 standard; protein; 502 AA.
AC ABU09594;
XX
DT 16-JUL-2003 (first entry)
DE Human cytochrome p450 gene CYP2D6, variant C5816AT, protein.
KW Human; cytochrome P450; CYP2D6; chromosome 22; SNP;
KW single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
KW psychiatric disorder; drug sensitivity.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 492..502
FT /note= "These amino acids replace the wild-type sequence
FT (Leu-Cys-Ala-Val-Pro-Arg) as the result of a single-
FT nucleotide polymorphism which causes a frameshift"
XX
PN EP1281755-A2.
XX
PD 05-FEB-2003.
XX
PF 16-JUL-2002; 2002BP-00254972.
XX
PR 31-JUL-2001; 2001US-0309111P.
XX (PF12) PFIZER PROD INC.
XX
XX Milos PM, Webb SM;
XX WPI; 2003-373769/36.
DR N-PSDB; ACA61304.
XX
XX New cytochrome p450 2D6 gene variants and polypeptides, useful for
PT determining if a subject has or is at risk of developing a drug
PT sensitivity condition or disorder that is associated with an aberrant
PT CYP2D6 activity.
XX
PS Claim 4; Fig 5; 88pp; English.

XX The invention relates to an isolated nucleic acid comprising a cytochrome
CC P450 2D6 gene variant, e.g. G5795C or C5816AT (referring to the genomic
CC sequence or the same variant nucleotide in the corresponding cDNA
CC sequences). Also included are probes, primers (allele specific
CC oligonucleotides) and arrays used to detect and/or amplify the CYP2D6
CC gene polymorphic regions, the variant polypeptides, antibodies which are
CC capable of distinguishing between the variant and wild-type polypeptides,
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
CC determining whether an individual is susceptible to being a poor
CC metaboliser of drugs. The DNA probe is useful for hybridising to a
CC variant form of the CYP2D6 gene. The primer is useful for amplifying the
CC C5816AT allelic variant. The allele specific nucleotide is useful for the
CC detection of the C5816AT allelic variant. The methods are useful for
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and
CC determining if an individual is susceptible to being a poor metaboliser
CC of drugs. The nucleic acids are useful as probes or primers for
CC determining whether a subject has a genetic deficiency for metabolising
CC drugs that are substrates of P450 CYP2D6. The methods are useful for
CC determining if a subject has or is at risk of developing a drug
CC sensitivity condition or disorder that is associated with an aberrant
CC CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
CC aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
CC appropriate drugs or determining the course of treatment to administer to
CC a subject to treat cardiovascular or psychiatric disorders, or for
CC treating a subject with a drug sensitivity or disorder associated with a
CC specific allelic variant of a polymorphic region of the CYP2D6 gene. The
CC antibodies are useful for monitoring CYP2D6 protein levels in an
CC individual for determining whether a subject has a disease or conditions
CC associated with an aberrant CYP2D6 protein level. The gene is located on
CC human chromosome 22. The present sequence is the C5816AT variant CYP2D6
CC protein
XX
SQ Sequence 502 AA;
Alignment Scores:
Pred. No.: 1.94e-164 Length: 502
Score: 2290.50 Matches: 447
Percent Similarity: 89.44% Conservative: 2
Best Local Similarity: 89.04% Mismatches: 2
Query Match: 79.37% Indels: 52
DB: 6 Gaps: 1
US-09-820-788A-1 (1-1537) x ABU09594 (1-502)
QY 78 ATGGGGCTAGAGCACTGGTGGCCCTGGCTGATAGTGGCCATCTTCTGCTCTGGTG 137
Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
QY 138 GACCTGTATGCACCGCGCCCAACGGCTGGCTGCACGCTACTCACAGGCCCTGCCACTG 197
Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProProGlyProLeuProLeu 40
QY 198 CCGCGGCTGGGCAACCTGCTGATGTGGACTTCCAGAACACACACCATCTGTTCCAGCAG 257
Db 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
QY 258 TTGCGGCGCGCTTCGGGAGCGTGTTCAGCTGACCTGAGCTGGCCCTGAGCGCGGTGCTG 317
Db 61 LeuArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
QY 318 CTCATATGGCTGGCGCGCGCTGGCGAGCGCTGGTGGACCCAGCGGAGACACCGCCGAC 377
Db 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
QY 378 CGCCCGCTGTGCCCCATCACCCAGATCTCTGGGCTTTTGGGCCCGCTTCCCAA----- 428
Db 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyValPhe 120
QY 428 ----- 428

Db 121 LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgPheSerValSerThrLeuArg 140
QY 428 ----- 428
Db 141 AsnLeuGlyLeuGlyLysSerLeuGluGlnTrpValThrGluGluAlaAlaCysLeu 160
QY 429 -----GGAGCCCTTTTCGCCCAACGGTCTCTTGACAAA 464
Db 161 CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys 180
QY 465 GCGGTGAGCAAGTATGCTCCCTCACCCTCGGGCGCGCTTCAGTACAGACCT 524
Db 181 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgPheGluTyrAspPro 200
QY 525 CCCTTCCTCAGGCTGCTGACCTAGCTCAGGAGGACTGAAGGAGGAGTGGCTTCGT 584
Db 201 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluLysSerGlyPheLeu 220
QY 585 CCGGAGGTGCTGAATGCTGTCCTCCCTCCTGCATATCCAGAGCTGCTGGCAAGGTC 644
Db 221 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 240
QY 645 CTACCTTCCTCCAAAGCTTTCTGACCCAGCTGATGATGCTTAACCTGAGCAGATG 704
Db 241 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuThrGluHisArgMet 260
QY 705 ACCTGGGACCCAGCCAGCCCGCCGAGACCTGACTGAGGCTTCCTGGCAGATGAG 764
Db 261 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
QY 765 AAGGCCAAGGGGAACCTCGAGAGCAGCTTCAATGATGAGAACCTCGCATAGTGTGCT 824
Db 281 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 300
QY 825 GACCTGTCTCTGCGGATGTGTACCACTCGACACCTGAGCTGGCGCTCTCTGCTC 884
Db 301 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeu 320
QY 885 ATGATCTTACATCCGATGTCAGCGCTGTCAGCGCTGTCACAGGAGATCGACGCTGATGG 944
Db 321 MetIleLeuHisProAspValGlnArgValGlnGlnGluIleAspValIleGly 340
QY 945 CAGGTGCGGCGACAGATGGTGACCAAGCTGACATGCTGCTCCCTGATTCACCACTGCTGAT 1004
Db 341 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrAlaValIle 360
QY 1005 CATGAGTGTGAGCGCTTTGGGAGCATCTGCTCCCTGGGTGTGACCCATATGACATCCCGT 1064
Db 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyMetThrHisMetThrSerArg 380
QY 1065 GACATCGAAGTACAGGGCTTCGGCATCCCTAAGGAACGACACTCATCACCACTGTCA 1124
Db 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400
QY 1125 TCGGTGCTCAAGGATGAGCGCTCTGGGAGAAGCCCTTCCTGCTCCACCCCGAACACTTC 1184
Db 401 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
QY 1185 CTGGATCCCGGCGGCACTTTGTGAAGCCGAGGCTTCCTGCTCTTCACGAGCGCGC 1244
Db 421 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 440
QY 1245 CGTGATCCCTCGGAGCCCTGCGCGCATGAGCTTCCTCTTCTTTCACCTCCCTG 1304
Db 441 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeu 460
QY 1305 CTGAGCACTTCAGCTTCTCGTGCCTGACAGCCCGCGCCAGCCACCATGGTGTCT 1364
Db 461 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 480
QY 1365 TTGCTTCTTCGTGACCCCAATCCCTATGAGCT-TTGTCTGTGCGCCCTAGATGG 1423
Db 481 PheAlaPheLeuValSerProSerProTyrGluTyrLeuCysCysAlaProLeuGluTrp 500

QY 1424 GGTACC 1429
|||||
Db 501 GlyThr 502

RESULT 5

AAR72376
ID AAR72376 standard; protein; 497 AA.
XX
AC AAR72376;
XX
DT 25-MAR-2003 (revised)
DT 15-NOV-1995 (first entry)
XX

Human auxillary cytochrome P450 species 2D6 variant 1 protein.

Human cytochrome P450; amplification; PCR; primer; expression vector;
yeast NADPH-P450 reductase; safety; fusion protein; metabolite;
carcinogen; mutagen; liver metabolism.

Homo sapiens.

Key Location/Qualifiers
Misc-difference 296
/note= "Cys to Arg variation"

EP644267-A2.

22-MAR-1995.

20-JUL-1994; 94EP-00111298.

20-JUL-1993; 93JP-00201120.

21-JUL-1993; 93JP-00180246.

30-JUL-1993; 93JP-00208279.

(SUMO) SUMITOMO CHEM CO LTD.

Hayashi K, Sakaki T, Yabusaki Y, Komai K, Kaneko H, Nakatsuka I;
WPI; 1995-116991/16.

N-PSDB; AAQ87730.
Evaluation of safety of a chemical cpd. - using recombinant yeast
expressing human cytochrome p450 and a yeast NADPH-P450 reductase.

Example; Page 87-89; 124pp; English.

The amino acid sequence of the human auxillary cytochrome P450 species
2D6 variant 1. This variant contains a variation at residue 296: Cys to
Arg, caused by a variation at base 886: T to C in the DNA sequence. The
cDNA was amplified by PCR using the primers AAQ87763-6. The product was
cloned into the yeast expression vectors pAAH5N or pAHR to produce the
vectors p2D6 variant 1 for the expression of the cytochrome P450 alone or
p2D6R variant 1 for co-expression with the yeast NADPH-P450 reductase.
The vectors are used in a method for evaluating the safety of a chemical
compound by reacting the chemical compound with recombinantly produced
human cytochrome P450 molecular species 1A2 (AAQ87714), 2C9 (AAQ87715),
2E1 (AAQ87716), or 3A4 (AAQ87718), or their auxillary species and
variants (AAQ87718-32), and yeast NADPH-P450 reductase, either as a fused
protein or in cell extracts, and analysing the resulting metabolite to
assess the safety of the chemical compound. The method is useful for
determining whether the chemical compound, or its metabolite, will be
converted into a carcinogenic or mutagenic form through metabolism in the
liver. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 497 AA;

Alignment Scores:

Pred. NO.: 2.3e-164 Length: 497
Score: 2289.50 Matches: 445
Percent Similarity: 89.54% Conservative: 0
Best Local Similarity: 89.54% Mismatches: 1

CC primers AAT26953-6. The prod. was cloned into the yeast expression vector
 CC pAAH5N to generate plasmid p2D6 for prodn. of the cytochrome only or into
 CC the vector pAHER to generate the plasmid p2D6R for co-prodn. with the
 CC yeast NADPH-P450 reductase. The sequence is placed under control of the
 CC yeast ADH gene promoter and terminator. The vectors are used in a method
 CC for evaluating the safety of a cpd. by reacting the test cpd. with
 CC recombinantly produced human cytochrome P450 mol. species 1A2 (AAT28380),
 CC 2C9 (AAT28381), 2E1 (AAT28382), 3A4 (AAT28383) or their variants
 CC (AAT28384-98) together with yeast NADPH-P450 reductase (either as a fused
 CC protein or as a cell extract) and analysing the resultant metabolite. The
 CC cpd. is considered "safe" if it is detoxified or not rendered
 CC carcinogenic or "unsafe" if it is not detoxified or is metabolised to a
 CC carcinogenic cpd
 XX
 SQ Sequence 497 AA;

Alignment Scores:

Pred. No.: 2,3e-164 Length: 497
 Score: 2289.50 Matches: 445
 Percent Similarity: 89.54% Conservative: 0
 Best Local Similarity: 89.54% Mismatches: 1
 Query Match: 79.33% Indels: 51
 DB: 2 Gaps: 1

US-09-820-788a-1 (1-1537) x AAR93183 (1-497)

QY 78 ATGGGGCTAGAGCACTGGTGGCCCTGGCGGTGATAGTGGCATCTTCTGCTCTGCTG 137
 DB 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
 QY 138 GACCTGATGACCGCGCGCAACGCTGGCTGACCTACTACACAGGCGCCCTGCCACTG 197
 DB 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaHisTyrProProGlyProLeuLeu 40
 QY 198 CCGGGCTGGCAACTGCTGCATGCTGCACTTCCAGAACACACCACTACTGCTTGCACGAG 257
 DB 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
 QY 258 TTGCGCGCGCGCTTCGGGACGTGTTACCTGACGCTGGCTGACCGCGGTGCTGCTG 317
 DB 61 LeuArgArgPheGlyAspValPheSerLeuGlnLeuAlaIleTrpProValVal 80
 QY 318 CTCATATGGCTGGCGCGCTGGCGCGCTGGTGACCCAGCGGAGGACACCGCGGAC 377
 DB 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
 QY 378 CGCCGCGCTGTGCCCATCACCCAGATCCTGGGTTTGGCGCGCTGCCAA----- 428
 DB 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyValPhe 120
 QY 428 ----- 428
 DB 121 LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgArgPheSerValSerThrLeuArg 140
 QY 428 ----- 428
 DB 141 AsnLeuGlyLeuGlyLysSerLeuGluGlnTrpValThrGluGluAlaAlaCysLeu 160
 QY 429 -----GGAGCGCCCTTTCCGCCCAACAGGTCTCTTGGACAAA 464
 DB 161 CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys 180
 QY 465 GCCGTGAGCAAGTATGCTCCCTCCCTACCTGGCGCGCGCTTCGAGTACGACACCT 524
 DB 181 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 200
 QY 525 CCGTCTCTCAGGCTGCTGACCTAGCTCAGGAGGAGCTGAAGGAGGAGTCCGGGCTTCTG 584
 DB 201 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 220
 QY 585 CCGGAGGTGCTGAATGCTGTGCTCCCGTCTCTGTCATATCCACGCGCTGGCGAAGTGC 644
 DB 221 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 240

QY 645 CTACGCTTCCAAAAGGCTTTCTCACCAGCTGGATGAGCTGCTAACTGAGCACAGGATG 704
 DB 241 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 260
 QY 705 ACCTGGGACCCAGACCCCGAGACCTGACTGAGGCTTCTCTGGCAGAGATGGAG 764
 DB 261 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
 QY 765 AAGCCAAAGGGGAACCTTGAGAGAGCTTCAATGATGAGAACCCTGCGCATGAGTGGCT 824
 DB 281 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 300
 QY 825 GACCTGTCTCTCGCGGATGGTGACCACTCGACCGCTGGCTGGCTGGGCTCTCTGCTC 884
 DB 301 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaIleTrpGlyLeuLeu 320
 QY 885 ATGATCTTACATCCGCGATGTGCAGCGCGCTGCCAACAGGAGATCGACGCTGATAGG 944
 DB 321 MetIleLeuHisProAspValGlnArgValGlnGlnGlnIleAspAspValIleGly 340
 QY 945 CAGGTGGCGGACACAGAGATGGGTGACAGGCTCACATGCCCTACACCTGCGCTGATT 1004
 DB 341 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 360
 QY 1005 CATGAGGTGACGCTTTGGGGACATCGTCCCTGGGTGTGACCCATATGACATCCCGT 1064
 DB 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 380
 QY 1065 GACATCGAAGTACAGGCTTCCGATCCCTAAGGGAACACACTCATCAACCACTGTCA 1124
 DB 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400
 QY 1125 TCGGTGTGAAGGATGAGCGCTCTGGAGAAAGCGCTTCCGCTTCCACCCGGAACACTTC 1184
 DB 401 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
 QY 1185 CTGATGATCCAGGCGCACTTTGTGAAGCGGAGCGCTTCTGCTGCTTCTCAGCAGGCGCG 1244
 DB 421 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuPhePheThrSerLeu 440
 QY 1245 CGTGATGCTTCGGGAGCGCCCTGGCGCGCATGGAGCTTCTCTCTTCTTCCCTCCCTG 1304
 DB 441 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 460
 QY 1305 CTGAGCACTTTCAGCTTCTCGGTGCGCACTGGACAGCGCGCGCGCAGCCACCATGCTGTC 1364
 DB 461 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 480
 QY 1365 TTTGCTTCTCTGCTGACCGCCATCCCTCATGAGCTTTGCTGCTGCTGCTGCTGCTGCTG 1415
 DB 481 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 497

RESULT 7

ADB25832
 ID ADB25832 standard; protein; 497 AA.
 XX
 AC ADB25832;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human CYP2D6-related protein #2.
 XX
 KW human; mutant CYP2D6 gene; drug analysis; drug testing.
 XX
 OS Homo sapiens.
 XX
 PN WO2003050282-A1.
 XX
 PD 19-JUN-2003.
 XX
 PF 05-DEC-2002; 2002WO-JP012748.
 XX

PR 06-DEC-2001; 2001JP-00372548.
XX (TSUR) TSUMURA & CO.
XX Taniyama M, Ogawa K, Teuchiya N, Hibino T;
XX WPI: 2003-505401/47.
XX N-PSDB; AD825777.
XX Genetic polymorphisms of CYP2D6 gene in human population for analysis of
XX drug effect on individual patients and testing of new drugs.
XX Claim 8; Page: 43-46; 75pp; Japanese.
XX The invention comprises mutant forms of the human CYP2D6 gene, containing
XX one or more of the following mutations G125A, C1858T, T2874C and C2875T.
XX The mutant human CYP2D6 genes of the invention are useful for analyzing
XX the effect of drugs on individual patients and testing of new drugs. The
XX present amino acid sequence represents a human protein of the invention.
XX Sequence 497 AA;
SQ
Alignment Scores:
Pred. No.: 2,74e-164 Length: 497
Score: 2288.50 Matches: 445
Percent Similarity: 89.54% Conservative: 0
Best Local Similarity: 89.54% Mismatches: 1
Query Match: 79.30% Indels: 51
DB: 6 Gaps: 1
US-09-820-788A-1 (1-1537) x ADB25832 (1-497)
QY 78 ATGGGGCTAGAGCACTGGTGGCCCTGGCGTATAGTGGCCATCTTCTGCTCTGCTG 137
DB 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
QY 138 GACCTGATCAGCGCGCAACGCTGGCTGACGCTACTACAGGCCCTGCGCCTG 197
DB 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40
QY 198 CCGGGCTGGCAACCTGCTGATGGACTTCCAGAACACACACTACTGCTTCAGCCAG 257
DB 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
QY 258 TTGGCGCGCGCTTGGGACGTGTTACGCTGCGAGCTGGCTGGACGCGGTGCTG 317
DB 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpProValValVal 80
QY 318 CTCATGGGCTGGCGCGCTGGCGAGCGCTGGTACCCAGCGGAGACACCGCGGAC 377
DB 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
QY 378 CCGCGCGCTGCGCCATCACCAGATCCCGGTTTGGCGCGCTGCCAA----- 428
DB 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyValPhe 120
QY 428 ----- 428
DB 121 LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgArgPheSerValSerThrLeuArg 140
QY 428 ----- 428
DB 141 AsnLeuGlyLeuGlyLysLysSerLeuGluGlnTrpValThrGluAlaAlaCysLeu 160
QY 429 -----GGAGCGCCCTTCCGCCCAACGGTCTCTGGACAAA 464
DB 161 CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys 180
QY 465 GCGGTGAGCAACGTGATCCCTCCCTCACCTCGCGCGCGCTTCCAGTACGACACCT 524
DB 181 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 200
QY 525 CGCTTCTCAGGCTGCTGGACCTAGCTCAGGAGGAGCTGAAGGAGGAGTTCGGCTTCTG 584

DB 201 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 220
QY 585 CGCAGAGTGTGAATGCTGCTCCCGTCTCTGTCATATCCAGCGCTGGCTGGCAAGGTC 644
DB 221 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyVal 240
QY 645 CTACGCTTCCAAAGGCTTCTGACCCAGCTGGATGAGCTGCTAACTGAGCAGGATG 704
DB 241 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 260
QY 705 ACCTGGGACCCAGCCCGACCCCGACCTGAGCTGAGGCTTCTTGGCAGAGATGGAG 764
DB 261 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
QY 765 AAGGCCAAGGGAACCTCGAGAGCTTCAATGATGAGACCTCGCATGCTGGTGGCT 824
DB 281 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 300
QY 825 GACCTGTTCTCTGCGGGATGGTGACCACTCGACCGCTGGCTGGCGCTCTCTGCTC 884
DB 301 AspLeuPheLeuAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 320
QY 885 ATGATCTTACATCCGGATGTCAGCGCGCTGTCCAAAGGAGATCGACGCTGATAGG 944
DB 321 MetIleLeuHisProAspValGlnArgValGlnGlnGluIleAspAspValIleGly 340
QY 945 CAGGTGCGCGCACAGAGATGGTGACCGCTCACTGCGCTACACCTACACCTGCGCTGAT 1004
DB 341 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrAlaValIle 360
QY 1005 CATGAGTGCAGCGCTTTGGGACATCGTCCCTGGGTGTGACCCATATGACATCCCGT 1064
DB 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 380
QY 1065 GACATCGAAGTACAGGCTTCCGCTCCCTAAGGGAACGACACTCATCACCACTGTCA 1124
DB 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400
QY 1125 TCGGTGTGAAGGATGAGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCGGAACATTC 1184
DB 401 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
QY 1185 CTGATGCCAGCGCCACTTTGTGAAGCCGAGGCGCTTCTGCGCTTCTTCAGCAGCGCGC 1244
DB 421 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 440
QY 1245 CGTGCATGCTCGGGGAGCCCTGCGCGCATGGAGCTTCTTCTTCTTACCTCCCTG 1304
DB 441 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 460
QY 1305 CTGACGACACTTCAGCTTCTCGGTGCGCCACTGACAGCCCGCGCCAGCCACCATGCTGTC 1364
DB 461 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 480
QY 1365 TTGCTGTTCTCTGAGCCCGCCATCCCGCTATGAGCTTGTGTGTGCGCCCGC 1415
DB 481 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 497
RESULT 8
ABU09598
ID ABU09598 standard; protein; 497 AA.
XX
AC ABU09598;
XX
DT 16-JUL-2003 (first entry)
XX
DE Human cytochrome p450 gene CYP2D6, variant G5799C, protein.
XX
KW Human: cytochrome P450; CYP2D6; chromosome 22; SNP;
KW single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
XX psychiatric disorder; drug sensitivity.
XX

QY	1005	CATGAGGTGCAGCGCTTTGGGGACATCGTCCCTCGGTGTGACCCATATGACATCCCGT	1064
Db	361	HisGluValGlnArgPheGlyAspIleValProLeuGlyMetThrHisMetThrSerArg	380
QY	1065	GACATCGAAGTACAGGGCTTCGGCATCCCTAAAGGAAGCACACTCATCCACCACTGTCA	1124
Db	381	AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer	400
QY	1125	TCGGTGTCTGAAGATGATGAGCGCTCTGGGAAAGCCCTTCGGCTTCCACCCGCAACACTTC	1184
Db	401	SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe	420
QY	1185	CTGGATCCCAAGGGCCACTTTGTGAAGCCGGAGGCCCTTCTTGCCTTTCTCAGCAGGCGCG	1244
Db	421	LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg	440
QY	1245	GTGTCATGCTCGGGAGCCCTTGGCCCGCATGTGAGCTCTTCCTCTCTTTCACCTCCCTG	1304
Db	441	ArgAlaCysLeuGlyGluProLeuAlaArgWetGluLeuPheLeuPheThrSerLeu	460
QY	1305	CTGCAGGCACTTCAGCTTCTCGTGTGCCCACTGGACAGCCCGCGCCACGCCACCATGTGTC	1364
Db	461	LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal	480
QY	1365	TTTGTCTTCTCGGTGACCCCATCCCCCTATGAGCTTTGTCTGTGCCCCCGC	1415
Db	481	PheAlaPheLeuValThrProSerProTrpGluLeuCysAlaValProArq97	

Score:	2286.50	Matches:	445
Percent Similarity:	89.54%	Conservative:	0
Best Local Similarity:	89.54%	Mismatches:	1
Query Match:	79.23%	Indels:	51
DB:	6	Gaps:	1
US-09-820-788A-1 (1-1537) x ADB25831 (1-497)			
QY	78	ATGGGGCTAGAAGCACTGGTGCCTCGCCGTGATAGTGGCCATCTTCTGCTCTCTGGTG	137
Db	1	MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal	20
QY	138	GACCTGATGACCCGGCCCAACGCTGGCTCAGCTACTCACCAGGCCCCCTGCCACTG	197
Db	21	AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu	40
QY	198	CCCCGGCTGGGCAACCTGCTCATGTGGACTTCAGAAACACACCACTACTGCTTCACACG	257
Db	41	ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln	60
QY	258	TTGCGGCGCCGCTTCGGGGACGTGTTCAGCTGCAGCTGGCTGGACCCGCTGGTG	317
Db	61	LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal	80
QY	318	CTCAATGGCTGGCGCGCTGCGGAGCGCTGTGTACGCTGCAGCTGGACCCAGGGAGGACACCCGCCGAC	377
Db	81	LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp	100
QY	378	CGCCCGCTGTGCCCATCACCCAGACTCTGGGTTTGGCGCGCTGCCA-----	428
Db	101	ArgProProValProIleThrGlnLeuGlyPheGlyProArgSerGlnGlyValPhe	120
QY	428	-----	428
Db	121	LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgArgPheSerValSerThrLeuArg	140
QY	428	-----	428
Db	141	AsnLeuGlyLeuGlyLysSerLeuGluGlnTrpValThrGluGluAlaAlaCysLeu	160
QY	429	-----GGACGCCCTTGC-----GGACGCCCTTGC-----GGACGCCCTTGC-----	464
Db	161	CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys	180
QY	465	GCCGTGAGCAACGTGATCGCTCCCTCACCCTCGCGGCGCCGCTTCAGCTAGCACCACTC	524
Db	181	AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro	200
QY	525	CGCTTCTCAGGCTGTGGACCTAGCTCAGAGGAGCTGAAGGAGGAGTCGGGCTTTCTG	584
Db	201	ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluLeuSerGlyPheLeu	220
QY	585	CGCAGGTGCTGAATGTGCTCCCGTCTCTGTGATATCCAGCGCTGGCTGGCAAGTTC	644
Db	221	ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal	240
QY	645	CTACGCTTCCAAAGGCTTCTCACCACCGCTGGATGAGCTGCTAACTGAGCACAGGATG	704
Db	241	LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet	260
QY	705	ACCTGGACCCAGCCCGCCCGAGACCTGACTGAGGCTTCTCCTGGCAGAGATGGAG	764
Db	261	ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu	280
QY	765	AAGCCCAAGGGGAACCTTGAGAGAGAGCTTCAATGATGAGAACCTCGCGCATAGTGTGCT	824
Db	281	LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuCysIleValValAla	300
QY	825	GACCTGTCTCTCGCGGATGGTGACCACTCGACCGCTGGCTGGCTGGGGCTCCTGCTC	884
Db	301	AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu	320
QY	885	ATGATCCTACATCCGGATGTGACGCGCGCTGTCCAAACAGGAGATCGACGCTGATAGG	944

QY 465 GCGGTGAGCAAGTGTATGCTCCCTCACCTGCGGGCGCGCTTCGAGTACGACACCT 524
DB 181 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgPheGluThrAspPro 200
QY 525 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGAGTCAAGAGAGTGGGCTTCG 584
DB 201 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluSerGlyPheLeu 220
QY 585 CGCAGGTGCTGAATGCTGCTCCCGCTCCTCGCATATCCAGCGCTGGCTGCAAGTGC 644
DB 221 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 240
QY 645 CTACGCTTCAAAGGCTTCTCGACCCAGCTGGATGAGTCTTAACGTGACACAGGATG 704
DB 241 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuThrGluHisArgMet 260
QY 705 ACTGGGACCCAGCCCGCCCGAGAGCTGAGTGGGCTTCTCGCAGAGTGGAG 764
DB 261 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
QY 765 AAGCCCAAGGGAACCTCGAGAGCAGCTTCAATGATGAGAACTGGCATAGTGGTGGCT 824
DB 281 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 300
QY 825 GACCTGTTCTCTCCCGGATGTCACCTCGACACCTGCGCTGGGCGCTCCTGCTC 884
DB 301 AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTrpGlyLeuLeuLeu 320
QY 885 ATGATCTACATCCGATCGAGCGCTGTCAGCGCTGTCACAGGAGTGCAGCTGTAGGG 944
DB 321 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspAspValIleGly 340
QY 945 CAGGTGCGCGACACGAGATGGTGACAGGCTTCATGCTCCCTGACCATCTCCCTGATT 1004
DB 341 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 360
QY 1005 CATGAGGTGACGCGCTTTCGGGACATCGTCCCGCTGGGTGTGACCCATATGACATCCCGT 1064
DB 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 380
QY 1065 GACATCGAAGTACAGGCTTCGCGATCCCTAGGGAGACGACACTCATACCACTGTCA 1124
DB 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400
QY 1125 TCGGTGCTCAAGATGAGCGCTGCGGAGAGCCCTTCGCTCCACCCGCAACACTTC 1184
DB 401 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
QY 1185 CTGGATGCCCGGCGCACCTTCTGAGCGGAGGCGCTTCTGCTTTCTCAGCAGCGCGC 1244
DB 421 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 440
QY 1245 CGTGATGCTCGGGAGCGCTGCGCGCGATGAGCTTCTCTCTTCTTCACTCCCTG 1304
DB 441 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrSerLeu 460
QY 1305 CTGCAGCACTTCAGCTTCTCGTGGCTGACGAGCGCGCGCGCGCGCGCGCGCGCG 1364
DB 461 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 480
QY 1365 TTTGCTTCTCCTGGTACCCCTCCCTATGAGCTTTGTGCTGTGCGCGCGCGCG 1415
DB 481 PheAlaPheLeuValSerProSerProTyrGluLeuCysAlaValProArg 497
RESULT 11
ID AAR93184
XX AAR93184 standard; protein; 497 AA.
AC AAR93184;
XX
XX 11-OCT-1996 (first entry)
DX

DE Human cytochrome P450 molecular species 2D6 variant #3 protein.
XX Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;
KW liver; Yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;
KW evaluation; safety; fusion protein; metabolite; detoxification;
XX carcinogenic.
OS Homo sapiens.
XX JP08056695-A.
XX 05-MAR-1996.
XX 15-JUL-1994; 94JP-00164184.
XX 20-JUL-1993; 93JP-00201120.
PR 30-JUL-1993; 93JP-00208279.
PR 17-JUN-1994; 94JP-00136053.
XX (SUMO) SUMITOMO CHEM CO LTD.
XX WPI; 1996-182311/19.
DR N-PSDB; AAT28397.
XX
PT Novel method for the evaluation of the safety of a cpd. - using a human
PT cytochrome P450 and yeast NADPH reductase to determine whether the
PT analyte cpd. is detoxified or metabolised to a carcinogen.
XX
PS Example 1; Page 53-55; 74pp; Japanese.
XX
CC This is the amino acid sequence of the human cytochrome P450 molecular
CC species 2D6 variant #3 protein. The corresp. gene was amplified from a
CC human liver derived cDNA library as 2 fragments of 0.4 and 0.9 kb using
CC primers AAT26953-6. The prod. was cloned into the yeast expression vector
CC pAAH5N to generate plasmid p2D6 for prodn. of the cytochrome only or into
CC the vector pAHRR to generate the plasmid p2D6R for co-prodn. with the
CC yeast NADPH-P450 reductase. The sequence is placed under control of the
CC yeast ADH gene promoter and terminator. The vectors are used in a method
CC for evaluating the safety of a cpd. by reacting the test cpd. with
CC recombinantly produced human cytochrome P450 mol. species 1A2 (AAT28380),
CC 2C9 (AAT28381), 2E1 (AAT28382), 3A4 (AAT28383) or their variants
CC (AAT28384-98) together with yeast NADPH-P450 reductase (either as a fused
CC protein or as a cell extract) and analysing the resultant metabolite. The
CC cpd. is considered "safe" if it is detoxified or not rendered
CC carcinogenic or "unsafe" if it is not detoxified or is metabolised to a
CC carcinogenic cpd
XX
SQ Sequence 497 AA;
Alignment Scores:
Pred. No.: 4,62e-164 Length: 497
Score: 2285.50 Matches: 444
Percent Similarity: 89.54% Conservative: 1
Best Local Similarity: 89.34% Mismatches: 1
Query Match: 79.19% Indels: 51
DB: 2 Gaps: 1
US-09-820-788A-1 (1-1537) x AAR93184 (1-497)
QY 78 ATGGGGCTAGAACGACCTGCTGCGCGCTGATGAGTGGCATCTTCTGCTCTGCTG 137
DB 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
QY 138 GACCTGATGACCGCGCGCAACGCTGCGCTGCGCTACTCACCAGGCCCTCGCCACTG 197
DB 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProProGlyProLeuLeu 40
QY 198 CCGGGCTGGGCAACCTGCTGCTGCTGACTTTCAGAAACACACACTACTGCTTCGACAC 257
DB 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
QY 258 TTGCGGCGCGCTTTCGGGAGCGTGTTCAGCTGCGAGCTGGCTGACCGCGCTGCTG 317

Db 61 LeuArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
QY 318 CTCATAGGCTGGCGCGCTGGCGAGCGCTGGTGCACCCGCGGAGGACACCGCCGAC 377
Db 81 LeuAenGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
QY 378 CGCCCGCTGTGCCCATCACCCAGATCTCTGGGTTTGGCGCGCTTCCCAA----- 428
Db 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyValPhe 120
QY 428 ----- 428
Db 121 LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgArgPheSerValSerThrLeuArg 140
QY 428 ----- 428
Db 141 AenLeuGlyLeuGlyLysSerLeuGluGlnTrpValThrGluGluAlaAlaCysLeu 160
QY 429 -----GGAGCGCCCTTTCGCCCAACCGGTCTCTTGGACAAA 464
Db 161 CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys 180
QY 465 GCCGTGAGCAACGTCATCCCTCTACCTGCGGCGCGCTTCGAGTACGACGACCT 524
Db 181 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 200
QY 525 CGCTTCTCAGGCTCTGACCTAGCTCAGAGGAGCTGAAGGAGGAGTGGGCTTCTG 584
Db 201 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 220
QY 585 CGCAGAGTGTGAATGCTGTCTCCCTCTCTGTCATATCCAGCGCTGGTGGCAAGGTC 644
Db 221 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 240
QY 645 CTACGCTTCCAAAGGCTTCTGACCCAGCTGGATGAGCTGCTAACTGAGCACAGGATG 704
Db 241 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 260
QY 705 ACCTGGGACCCAGCCCGCCCGAGACCTGACTGAGGCTTCTCGCAGAGATGAG 764
Db 261 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
QY 765 AAGGCCAAGGGGAACCTCGAGAGCAGCTTCAATGATGAGAACCTCGCATAGTGTGCT 824
Db 281 LysAlaLysGlyAsnProGluSerPheAsnAspGluAsnLeuArgIleValValAla 300
QY 825 GACCTGTCTCTGCGCGATGTGACCACTGACCGCTGCGCTGGCGCTCTGCTC 884
Db 301 AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTrpGlyLeuLeuLeu 320
QY 885 ATGATCTTACATCCGATGTGCGCGCTGTCACACAGGAGATCGACGCTGATAGG 944
Db 321 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspAspValIleGly 340
QY 945 CAGGTGCGCGCAGACAGATGGTGACCAAGGCTCATGCGCTCACACCACTGCGCTGATT 1004
Db 341 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrAlaValIle 360
QY 1005 CATGAGTGCAGCGCTTTGGGACATCGTCCCTGGGTGTGACCCATATGACATCCCGT 1064
Db 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 380
QY 1065 GACATCGAAGTACAGGGCTTCGGCATCTCTAAGGAAGACACTCATCACCACTGTCA 1124
Db 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400
QY 1125 TCGGTGCTGAAGGATCAGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCCGAACACTTC 1184
Db 401 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
QY 1185 CTGGATGCCAGGGCCATTTGTGAAGCGGAGGCTTCTGCTGCTTTCTCAGCAGCGCGC 1244
Db 421 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 440

QY 1245 CGTCATGCTCGCGGAGCCCTGGCCGCGATGAGCTCTTCTTCTTCCACCTCCCTG 1304
Db 441 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeu 460
QY 1305 CTCGAGCATTTCAGCTTCTCGGTGCCCACTGAGCAGCCCGCCAGCCACCATGGTGTG 1364
Db 461 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 480
QY 1365 TTTCGTTCTCTGGTGACCCCATCCCTATGAGCTTTGTGCTGTGCCCGC 1415
Db 481 PheAlaPheLeuValSerProSerProTyrGluLeuCysAlaValProArg 497

RESULT 12

AAR81462
ID AAR81462 standard; protein; 497 AA.

AC AAR81462;

XX 01-AUG-1996 (first entry)

XX Human derived cytochrome P4502D6.

XX Human derived cytochrome; P4502D6; commercial cDNA library; yeast;
transfection; recombinant production; expression vector; mammal;
KW immunisation; sensitisation; antibody; determination; detection;
KW non-cross reactive.

XX Homo sapiens.

OS Homo sapiens.

PN JP08027199-A.

XX 30-JAN-1996.

XX 15-JUL-1994; 94JP-00164186.

XX 15-JUL-1994; 94JP-00164186.

XX (SUMO) SUMITOMO CHEM CO LTD.

XX WPI; 1996-136340/14.

DR N-PSDB; AAT17388.

XX Antibody recognising human derived cytochrome P4502D6 - allows specific
detection of cytochrome P450 species in humans.

XX Example 1; Page 11-13; 13pp; Japanese.

XX The present sequence is the human derived cytochrome (HDC) P4502D6, which
was obtd. from a commercial cDNA library. Yeast were transfected with an
expression vector contg. the HDC cDNA, cultured and then disrupted to
give a microsomal fraction. The HDC was purified from the fraction, and
used to immunise and sensitise a mammal. Blood was drawn from the mammal,
and an anti-HDC antibody isolated. The antibody obtd. recognises HDC
P4502D6, partic. at a serum dilution rate of 1:10000, and is
substantially without cross reaction to other HDC P450 spp

XX Sequence 497 AA;

Alignment Scores:

Pred. No.: 4.62e-164 Length: 497

Score: 2285.50 Matches: 444

Percent Similarity: 89.54% Conservative: 1

Best Local Similarity: 89.34% Mismatches: 1

Query Match: 79.19% Indels: 51

DB: 2 Gaps: 1

US-09-820-788A-1 (1-1537) x AAR81462 (1-497)

QY 78 ATGGGGCTAGAACACTGGTGGTGGCCCTGGCCCTGATGGCCATCTTCTCTCTGCTG 137

Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20

138 GACCTGATGACCGGCGCCAAACCTGGCTGGCGTACGCTACTACACAGGCCCTCGCACCTG 197
Db |||||LeuMetHisArgArgGlnArgTrpAlaAlaArgTrpProGlyProLeuProLeu 40
198 CCCGGGCTGGGCAACCTGCTGCATGTGGACTCCAGAACACACACATACCTGCTCGACCCAG 257
Db |||||GlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
258 TTCCGGCGCGCTTCGGGACAGTGTTCAGCCTGCAGCTGGCTCGGACCGCGGTGGTGGT 317
Db |||||LeuArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValVal 80
318 CTCATATGGCTGGCGCGCTGGCGAGCGCTGGTGAACCCAGCAGACACCGCCGAC 377
Db |||||LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyLeuAspThrAlaAsp 100
378 CGCGCCGCTGTCGCCATCCACAGATCCCTGGGTTCGGCGCGGTCCCAA----- 428
Db |||||ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyValPhe 120
428 ----- 428
Db |||||LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgPheSerValSerThrLeuArg 140
428 ----- 428
Db |||||AsnLeuGlyLeuGlyLysLysSerLeuGluGlnTrpValThrGluGluAlaAlaCysLeu 160
429 -----GGACGCGCTTCGGCCCAACCGTCTCTTGGACAA 464
Db |||||CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys 180
465 GCGGTGAGCAACGTGATCGCTCCCTCACCTGGCGCGCGCTTCAGTACGACGACCT 524
Db |||||AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTrpAspAspPro 200
525 CGCTTCCTCAGGCTGTGACCTAGCTCAGGAGGACTGAAGAGGAGTGGGCTTCG 584
Db |||||ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 220
585 CGCGAGGTGCTGATGCTGCTCCGCTCCCTGCTGCATATCCAGCGCTGCTGGCAGGTC 644
Db |||||ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 240
645 CTACGCTTCCAAAGGCTTCCTGACCCAGCTGATGAGTGTCTAAGTACGACACAGGATG 704
Db |||||LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 260
705 ACCTGGGACCCAGCCCGCCCGAGACCTGACTGAGGCGCTTCCTGGCAGAGATGGAG 764
Db |||||ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
765 AAGCCCAAGGGAAACCTGAGAGCAGCTTCAATGATGAGAACCTGGCAGTATGTTGGCT 824
Db |||||LysAlaLysGlyAsnProGluSerPheAsnAspGluAsnLeuArgIleValValAla 300
825 GACCTGTTCTCTCGGGATGTGACCACTGACCGCTGCTGCTGCTGGCGCTCTGCTC 884
Db |||||AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 320
885 ATGATCTACATCCGATGTGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944
Db |||||MetIleLeuHisProAspValGlnArgArgValGlnGlnGlnIleAspAspValIleGly 340
945 CAGGTGCGCGGACGAGATGGGTGACCAAGGCTCACATGCTGCTGCTGCTGCTGCTGCTGCT 1004
Db |||||GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 360
1005 CATGAGTTCAGGCTTCGGGACATCGTCCCGCTGGGTGTGACCATATGATCATCCCGT 1064
Db |||||HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 380
1065 GACATCGAAGTACAGGGCTTCGCGATCCCTTAAGGGAAACGACACTCATCAACCACTGTCA 1124

381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400
1125 TCGGTGCTCAAGSANTAGCGCTCTGGGAGAGCCCTTCGGCTTCCACCCCGAACACTTC 1184
Db |||||SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
1185 CTGGATGCCAGCGGCACCTTTGTGAAGCGGAGCGCTTCCTGCTTTCTCAGCAGCGCCG 1244
Db |||||LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 440
1245 CGTGATGCTCGGGAGCGCTTCGGCCCGCATGGAGCTTCTCTTCTTCTCCTCCTG 1304
Db |||||ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 460
1305 CTGCAGCACTTCAGCTTCTCGTGGCCCTGACGACGCGCGCCGCGCCAGCCACCATGCTGTC 1364
Db |||||LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 480
1365 TTTGCTTTCTGCTGACCCATCCCTATGAGCTTTGTGCTGTGCGCCGCGC 1415
Db |||||PheAlaPheLeuValSerProSerProTyrGluLeuCysAlaValProArg 497
RESULT 13
ABB09563
ID ABB09563 standard; protein; 497 AA.
XX
AC ABB09563;
XX
DT 02-SEP-2002 (first entry)
XX
DE Human CYP2D6 protein, SEQ ID NO:3.
XX
KW Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;
chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;
antiarrhythmic; arrhythmia; adrenoceptor antagonist; hypertension;
tricyclic antidepressant; procainamide; drug induced lupus syndrome;
environmentally linked disease; Parkinson's disease; haplotyping;
genotyping; haplotype; genetic variant; single nucleotide polymorphism;
SNP; drug screening; drug discovery.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 7 /note= "Encoded by RTG in the sequence given in ABQ72215.
This residue is Met (encoded by ATG) rather than Val
(encoded by GTG) in a polymorphic variant"
FT FT Misc-difference 11 /note= "Encoded by RTG in the sequence given in ABQ72215.
This residue is Met (encoded by ATG) rather than Val
(encoded by GTG) in a polymorphic variant"
FT FT Misc-difference 34 /note= "Encoded by YCA in the sequence given in ABQ72215.
This residue is Ser (encoded by TCA) rather than Pro
(encoded by CCA) in a polymorphic variant"
FT FT Misc-difference 88 /note= "Encoded by CRC in the sequence given in ABQ72215.
This residue is His (encoded by CAC) rather than Arg
(encoded by CGC) in a polymorphic variant"
FT FT Misc-difference 91 /note= "Encoded by MTG in the sequence given in ABQ72215.
This residue is Met (encoded by ATG) rather than Leu
(encoded by CTG) in a polymorphic variant"
FT FT Misc-difference 94 /note= "Encoded by CRC in the sequence given in ABQ72215.
This residue is Arg (encoded by GCG) rather than His
(encoded by CAC) in a polymorphic variant"
FT FT Misc-difference 98 /note= "Encoded by ACS in the sequence given in ABQ72215"
FT FT Misc-difference 104 /note= "Encoded by GYG in the sequence given in ABQ72215.
This residue is Ala (encoded by GCG) rather than Val

FT Misc-difference 107 (encoded by GTG) in a polymorphic variant"
FT /note= "Encoded by WTC in the sequence given in ABQ72215.
FT This residue is Phe (encoded by TTC) rather than Thr
FT (encoded by ACC) in a polymorphic variant"
FT Misc-difference 109
FT /note= "Encoded by RTC in the sequence given in ABQ72215.
FT This residue is Val (encoded by GTC) rather than Ile
FT (encoded by ATC) in a polymorphic variant"
FT Misc-difference 111
FT /note= "Encoded by GGY in the sequence given in ABQ72215"
FT /note= "Encoded by TTY in the sequence given in ABQ72215"
FT Misc-difference 120
FT /note= "Encoded by WTC in the sequence given in ABQ72215.
FT This residue is Ile (encoded by ATC) rather than Phe
FT (encoded by TTC) in a polymorphic variant"
FT Misc-difference 128
FT /note= "Encoded by YGG in the sequence given in ABQ72215.
FT This residue is Arg (encoded by CGG) rather than Trp
FT (encoded by TGG) in a polymorphic variant"
FT Misc-difference 136
FT /note= "Encoded by RTS in the sequence given in ABQ72215.
FT This residue is Ile (encoded by ATC) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT Misc-difference 151
FT /note= "Encoded by SAG in the sequence given in ABQ72215.
FT This residue is Glu (encoded by GAG) rather than Gln
FT (encoded by CAG) in a polymorphic variant"
FT Misc-difference 155
FT /note= "Encoded by RAG in the sequence given in ABQ72215.
FT This residue is Lys (encoded by AAG) rather than Glu
FT (encoded by GAG) in a polymorphic variant"
FT Misc-difference 232
FT /note= "Encoded by CAY in the sequence given in ABQ72215"
FT /note= "Encoded by RTC in the sequence given in ABQ72215.
FT This residue is Met (encoded by ATG) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT Misc-difference 344
FT /note= "Encoded by YGA in the sequence given in ABQ72215.
FT This is a stop codon (TGA) rather than Arg (encoded by
FT CGA) in a polymorphic variant"
FT Misc-difference 361
FT /note= "Encoded by CAY in the sequence given in ABQ72215"
FT /note= "Encoded by ASC in the sequence given in ABQ72215.
FT This residue is Thr (encoded by ACC) rather than Ser
FT (encoded by AGC) in a polymorphic variant"
FT Misc-difference 497
FT /note= "Encoded by YGC in the sequence given in ABQ72215.
FT This residue is Cys (encoded by TGC) rather than Arg
FT (encoded by CGC) in a polymorphic variant"
XX
XX WO200238589-A2.
XX
XX 16-MAY-2002.
XX
XX 09-NOV-2001; 2001WO-US047396.
XX
XX 09-NOV-2000; 2000US-0247943P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Anastasio AE, Chew A, Choi JY, Denton RR, Nandabalan K;
XX Petersen N, Rounds E;
XX
XX WPI; 2002-519292/55.
XX
XX N-PSDB; ABQ72215, ABQ72216, ABQ72364.
XX
XX Novel genetic variants of Cytochrome P450, Subfamily IID, Polypeptide 6
XX isoenzymes, useful for improving efficiency and reliability in drug
XX development for treating hypertension, arrhythmias and Parkinson's

PT disease.
XX Claim 29; Fig 3; 158pp; English.
XX
XX The invention relates to a method for haplotyping the cytochrome P450,
XX subfamily IID, polypeptide 6 (CYP2D6) gene (ABQ72215, ABQ72364) of an
XX individual, and also describes 29 novel polymorphic sites within the
XX human CYP2D6 gene. The CYP2D6 gene is located on chromosome 22q13.1 and
XX contains 9 exons which encode a 497 amino acid protein (ABB09563). CYP2D6
XX is a mono-oxygenase involved in the detoxification of many drugs and
XX environmental chemicals. It plays a role in the metabolism of drugs such
XX as antiarrhythmics, adrenoceptor antagonists and tricyclic
XX antidepressants, and is also involved in the formation of a metabolite
XX linked to the drug-induced lupus syndrome observed with procainamide.
XX Variations in CYP2D6 activity or expression may also influence an
XX individual's susceptibility to environmentally-linked diseases, and it
XX has been demonstrated that CYP2D6 activity may be involved in the
XX pathogenesis of Parkinson's disease, with individuals with a less active
XX form of the enzyme tending to have an earlier onset of this condition.
XX CYP2D6 nucleic acid sequences are useful in studying the expression and
XX function of CYP2D6, and in expressing CYP2D6 protein for use in screening
XX drugs for the treatment of CYP2D6-associated diseases (e.g.,
XX hypertension, atrial and ventricular arrhythmias, Parkinson's disease,
XX and drug-induced lupus syndrome) or which are metabolised by CYP2D6.
XX CYP2D6 nucleic acids and proteins are also useful in studying the effect
XX of polymorphisms on the biological activity of CYP2D6. Polymorphisms in
XX the target region may be determined by the use of allele-specific
XX oligonucleotides (ASOs; ABQ72217-ABQ72303) as probes and primers, and by
XX primer extension using oligonucleotide primers comprising sequences
XX ABQ72304-ABQ72361. The method of the invention is useful for haplotyping
XX the CYP2D6 gene in populations and in individuals, enabling decisions to
XX be made as to whether CYP2D6 is a likely therapeutic target for a disease
XX of interest, and to control for genetically-based bias in the design of
XX drugs that target or are metabolised by CYP2D6. In addition, transgenic
XX animals comprising a human CYP2D6 gene are useful for studying the
XX expression of CYP2D6 isoenzymes in vivo, for in vivo screening and testing
XX of drugs targeted to or metabolised by CYP2D6, and for testing the
XX efficacy of therapeutic agents and compounds for treating CYP2D6-
XX associated conditions in a biological system. The present sequence
XX represents the specifically claimed human CYP2D6 protein. This sequence
XX contains 18 polymorphic sites caused by polymorphisms in the coding
XX sequence (ABQ72216)
XX
XX Sequence 497 AA;
XX
XX Alignment Scores:
XX Pred. No.: 4.62e-164 Length: 497
XX Score: 2285.50 Matches: 444
XX Percent Similarity: 89.54% Conservative: 1
XX Best Local Similarity: 79.34% Mismatches: 1
XX Query Match: 79.19% Indels: 51
XX DB: 5 Gaps: 1
XX
XX US-09-820-788A-1 (1-1537) x ABB09563 (1-497)
XX
XX QY 78 ATGGGGCTAGAGCACTGGTGGCCCTGGCGCTAGTGGCCATCTTCTCTCTCTGTG 137
XX DB 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
XX
XX QY 138 GACCTGATGCACCGCGCCAAACGGCTGGGCTGCAGCTACTCACCAGGCCCTCGCACTG 197
XX DB 21 AspleuMetHisArgArgGlnArgTrpAlaalaargTyProProGlyProLeuProLeu 40
XX
XX QY 198 CCGGGCTGGGCAACCTGCTCATGTGAGCTTCAGAAACACACCATATCTGCTTCACGAG 257
XX DB 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAenThrProTyCysPheAspGln 60
XX
XX QY 258 TTGGGGCCCGCTTGGGGAGCTGTTCAGCTGCGAGCTGGCGTGGACCCCGCTGTGCTGTG 317
XX DB 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaIleTrpProValValVal 80
XX
XX QY 318 CTCAATGGCTGGCGCGCGTGGCGAGCGCTGTGACCCACGGCGAGGACACCCCGGAC 377
XX

Db 81 LeuAsnGlyLeuAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
QY 378 CGCCCGCTGTGCCCCATCCAGATCCTGGGTTTGGGCGCGTTCCTCCAA 428
Db 101 ArgProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyValPhe 120
QY 428 ----- 428
Db 121 LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgPheSerValSerThrLeuArg 140
QY 428 ----- 428
Db 141 AsnLeuGlyLeuGlyLysSerLeuGluGlnTrpValThrGluAlaAlaCysLeu 160
QY 429 -----GGAGCCCGCTTTCCGCCCAACCGTCTCTTTGGACAAA 464
Db 161 CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys 180
QY 465 GCGCTGAGCAACGTGATCGCTCCCTCACCTGCGGGCGCGCTTCGAGTACGACGACCT 524
Db 181 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgPheGluTyrAspAspPro 200
QY 525 CGCTTCCTCAGGCTGTGACCTAGCTCAGGAGGACTGAAGGAGAGTCGGGCTTCG 584
Db 201 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluSerGlyPheLeu 220
QY 585 CGCAGGTGCTGAATCTGTCCCGCTCTCCATATCCAGCGCTGCTGCAAGTTC 644
Db 221 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 240
QY 645 CTACGCTTCAAAGCTTTCCTGACCCAGCTGGATGAGTCTAACTGACGACAGGATG 704
Db 241 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuThrGluHisArgMet 260
QY 705 ACTGCGGACCCAGCCCGCCCGAGACCTGACTGAGCGCTTCTGCGCAGAGTGGAG 764
Db 261 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
QY 765 AAGGCCAAGGGGAACCTCAGAGCAGCTTCAATGATGAGAACTGCGCATGTGGTGGCT 824
Db 281 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValAla 300
QY 825 GACCTGTTCTGCGCGGATGTGACCACTCGACACCGCTGCGCTGGCGCTCTCTGCTC 884
Db 301 AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTrpGlyLeuLeuLeu 320
QY 885 ATGATCTCATCCGATGTGACGCGCTGTCCACAGAGATCGACACGATGATAGG 944
Db 321 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspAspValIleGly 340
QY 945 CAGGTGCGCGGACACGAGATGGTGACAGGCTCACATGCTTACACCACTGCGCTGATT 1004
Db 341 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 360
QY 1005 CATGAGGTGACGCGCTTTCGGGACATCGTCCCGCTGGGTGTGACCCATATGACATCCCGT 1064
Db 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 380
QY 1065 GACATCGAAGTACAGGCTTCGCGATCCCTAAGGGAAACGACACTCATCAACCACTGTCA 1124
Db 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400
QY 1125 TCGGTGCTCAAGGATGAGCGCTCTGGGAGAGCCCTTCGGTTCACCCCGCAACACTTC 1184
Db 401 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
QY 1185 CTGGATGCCAGGGCCACTTTGTGAAGCCGAGGCGCTTCTGCTTTCTCAGCAGCGCGC 1244
Db 421 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 440
QY 1245 CTGTCATGCTCGGGAGCCCTGCGCGCATGGAGCTTCTCTCTTTCACCTCCCTG 1304
Db 441 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 460

QY 1305 CTGACGACTTTCAGCTTCTCGGTGCCCTGACAGACCCCGCCCGCCACCATGGTGTTC 1364
Db 461 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 480
QY 1365 TTTGCTTTCTCGTGGTACCCCATCCCGCTATGAGCTTTGCTGTGCCCCGC 1415
Db 481 PheAlaPheLeuValSerProSerProTyrGluLeuCysAlaValProArg 497

RESULT 14

AAO26405
ID AAO26405 standard; protein; 497 AA.

AC AAO26405;

DT 30-JAN-2003 (first entry)

DE Human drug-metabolising enzyme related protein.

XX Cytostatic; antiparkinsonian; gene therapy; cancer; Parkinson's disease;
human drug-metabolising protein; enzyme.

OS Homo sapiens.

XX WO200279233-A1.

PN 10-OCT-2002.

XX 01-APR-2002; 2002WO-US009738.

XX 30-MAR-2001; 2001US-00820788.

XX (PEKE) PE CORP NY.

PA (DPEA/) DI FRANCESCO V.

PA (BEAS/) BEASLEY E M.

XX Shao W, Yan C;

XX WPI; 2003-040649/03.

XX New human drug-metabolizing proteins and nucleic acids related to the
Cytochrome P450 IID drug-metabolizing enzyme subfamily, useful for
treating a condition mediated by a human enzyme protein e.g., cancer.

XX Disclosure; Fig 2B; 72pp; English.

XX The invention relates to a novel isolated polypeptide comprising a 446-
amino acid sequence or its allelic variant, orthologue or fragment. The
allelic variant or orthologue is encoded by a nucleic acid that
hybridises under stringent conditions to the opposite strand of the
nucleic acid comprising a sequence having 1537 or 10278 bp. The fragment
comprises at least 10 contiguous amino acids of the 446-amino acid
sequence. The polypeptide is useful for preparing a pharmaceutical
composition for treating a disease or condition mediated by a human
enzyme protein, e.g. cancer or Parkinson's disease. This sequence
represents a human drug-metabolising related protein of the invention

SQ Sequence 497 AA;

Alignment Scores:

Pred. No.:	4,628-164	Length:	497
Score:	2285.50	Matches:	444
Percent Similarity:	89.54%	Conservative:	1
Best Local Similarity:	89.34%	Mismatches:	1
Query Match:	79.19%	Indels:	51
DB:	6	Gaps:	1

US-09-820-788A-1 (1-1537) x AAO26405 (1-497)

QY 78 ATGGGGCTAGAACGACTGGTCCCGCTGAGTAGTGGCCATCTTCTGCTCTCTGGTG 137

Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20

```
QY 138 GACGTGATGACCGCGGCAACGCTGGCTGACCTACTACACAGGCCCTCCACTG 197
Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProProGlyProLeuProLeu 40
QY 198 CCCGGCTGGCAACCTCTGCTGATCTGAGCTTCCAGAACACACACATGCTTCCAGCAG 257
Db 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
QY 258 TTGCGCGCGCGCTTCCGGGACGCTGTTACGCTTGACGCTGGCTGACGCGCGTGGCTGCTG 317
Db 61 LeuArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
QY 318 CTCATAGGCTGGCGCGCTGCGCGAGCGCTGTGACCCACGCGAGAGACACCGCCGAC 377
Db 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
QY 378 CGCCCGCTGTGCTCCATCACCAGATCTGGGTGTTGGCGCGCTTCCCAA----- 428
Db 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyValPhe 120
QY 428 ----- 428
Db 121 LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgArgPheSerValSerThrLeuArg 140
QY 428 ----- 428
Db 141 AsnLeuGlyLeuGlyLysLysSerLeuGluGlnTrpValThrGluGluAlaAlaCysLeu 160
QY 429 -----GGAGCGCCCTTTCGCCCGCAACGGTCTCTTGACAAA 464
Db 161 CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys 180
QY 465 GCGGTGAGCAACGTGATGCTGCTCCCTCCCTGCGGCGCGCTTCCAGTACGACGACCT 524
Db 181 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluThrAspAspPro 200
QY 525 CGCTTCTCAGGCTGCTGACCTAGCTCAGAGGGGACTGAAGGAGGAGTCCGGCTTCTG 584
Db 201 ArgPheLeuArgLeuLeuAspLeuAlaGlnGlyLeuLysGluGluSerGlyPheLeu 220
QY 585 CGCAGGTGCTGAATGCTGTCCCTCTCTGCTGATATCCAGCGCTGCTGGCGAGTGC 644
Db 221 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 240
QY 645 CTACGCTTCCAAAGGCTTCTCTGACCCAGCTGAGTGTGCTAACTGAGCAGCAGGATG 704
Db 241 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 260
QY 705 ACCTGGGACCCAGCCCGCCCGAGACCTGAGGCTTCTCTGCGCAGAGATGGAG 764
Db 261 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
QY 765 AAGGCCAGGGAAACCTGAGAGAGCTTCAATGATGAGACCTCGCATAGTGTGCT 824
Db 281 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 300
QY 825 GACCTGTTCTCTCGCGGATGTGACACCTCGACCGCTGGCTGGCGCTCTGCTC 884
Db 301 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 320
QY 885 ATGATCTACATCCGAGTGTGAGCGCGCTGTCCAAACAGGAGATCGACGACGTGATAGG 944
Db 321 MetIleLeuHisProAspValGlnArgValGlnGlnIleAspAspValIleGly 340
QY 945 CAGGTGCGCGGACAGAGATGGGTACAGGCTACATGCCCTACACACTGCCGTGATT 1004
Db 341 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 360
QY 1005 CATGAGTGCAGCGCTTTGGGACATCGTCCCTGGGTGTGACCCATATGACATCCCGT 1064
Db 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 380
QY 1065 GACATCGAAGTACAGGGCTTCCGCATCCCTAAGGGAGACGACACTCATCACCAACCTGTCA 1124
```

```
Db 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400
QY 1125 TCGGTGCTGAAGGATAGGCGCTCTGGGAGAGCCCTTCGCTTCCACCCGAAACACTTC 1184
Db 401 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
QY 1185 CTGATGCCAGGCGCACTTTGTGAAGCGGAGGCTTCTCCCTTTCTCAGCAGCGCGC 1244
Db 421 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 440
QY 1245 CGTGCATGCTCGGGAGCGCCCTGCGCGGATGAGAGCTTCTTCTTCTCCTCCCTG 1304
Db 441 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 460
QY 1305 CTGCAGGACCTTACGCTTCTCGTGTCGCCACTCGACAGCCCGCGCAGCACCACATGCTC 1364
Db 461 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 480
QY 1365 TTTGCTTTCTGCTGACCCCATCCCTATGAGCTTTGTGCTGTGCCCGC 1415
Db 481 PheAlaPheLeuValSerProSerProTyrGluLeuCysAlaValProArg 497

RESULT 15
ADB25833
ID ADB25833 standard; protein; 497 AA.
XX
AC ADB25833;
XX
DT 20-NOV-2003 (first entry)
DE Human CYP2D6-related protein #3.
XX
KW human; mutant CYP2D6 gene; drug analysis; drug testing.
XX
OS Homo sapiens.
XX
PN WO2003050282-A1.
XX
PD 19-JUN-2003.
XX
PF 05-DEC-2002; 2002WO-JP012748.
XX
PR 06-DEC-2001; 2001JP-00372548.
XX
PA (TSUR ) TSUMURA & CO.
XX
PI Taniyama M, Ogawa K, Tsuchiya N, Hibino T;
XX
WPI; 2003-505401/47.
DR N-PSDB; ADB25778.
XX
Genetic polymorphisms of CYP2D6 gene in human population for analysis of
PT drug effect on individual patients and testing of new drugs.
XX
PS Claim 8; Page 46-50; 75pp; Japanese.
XX
CC The invention comprises mutant forms of the human CYP2D6 gene, containing
CC one or more of the following mutations G125A, C1858T, T2874C and C2875T.
CC The mutant human CYP2D6 genes of the invention are useful for analysing
CC the effect of drugs on individual patients and testing of new drugs. The
CC present amino acid sequence represents a human protein of the invention.
XX
SQ Sequence 497 AA;

Alignment Scores:
Pred. No.: 4,62e-164 Length: 497
Score: 2285.50 Matches: 444
Percent Similarity: 89.54% Conservative: 1
Best Local Similarity: 89.34% Mismatches: 1
Query Match: 79.19% Indels: 51
DB: 6
```

US-09-820-788A-1 (1-1537) x ADB25833 (1-497)

```
QY 78 ATGGGGCTAGAAGCACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCTGCTCTCTGGTG 137
Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
QY 138 GACCTGATGACCGGGCCCAACGCTGGCTGCGCTACTACTACAGGCCCTCCGCACTG 197
Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProProGlyProLeuProLeu 40
QY 198 CCGGGGCTGGGCAACCTGTGCTATGGGCTTCCAGAACACACACACTACTCTCGACCA 257
Db 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
QY 258 TTGCGCGCCGCTTCGGGACCGTGTTCAGCCTGCGCTGGCGCTGGACCGCGGTGCTG 317
Db 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
QY 318 CTCATGGGCTGGCGGCGCTGGCGAGGCGTGGTACCCACGCGGAGGACACCGCGGAC 377
Db 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
QY 378 CGCGCGCTGTGCCATACCCAGATCCTGGGTGTTGGCGCGCTGCCAA----- 428
Db 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyValPhe 120
QY 428 ----- 428
Db 121 LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgArgPheSerValSerThrLeuArg 140
QY 428 ----- 428
Db 141 AsnLeuGlyLeuGlyLysLysSerLeuGluGlnTrpValThrGluGluAlaAlaCysLeu 160
QY 429 -----GGACGCCCTTTGCGCCCAACGGTCTCTGGACAAA 464
Db 161 CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys 180
QY 465 GCCGTGAGCAAGTGATCGCTCCCTCACCTCGGGCGCGCTTCGAGTACGACGACCT 524
Db 181 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 200
QY 525 CGCTTCTCAGGCTGTGACCTAGCTCAGGAGGAGCTGAAGGAGGAGTCGGGCTTCTG 584
Db 201 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 220
QY 585 CGCGAGGTCTGAATGCTCTCCCTGCTCTGCTATCCAGCGCTGGCTGGCAAGTTC 644
Db 221 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 240
QY 645 CTACGCTTCCAAAAGGCTTCTGACCCAGCTGGATGAGCTGCTAACTGACGACAGGATG 704
Db 241 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 260
QY 705 ACCTGGGACCCAGCCCGCCGAGACCTGACTGAGGCTTCTGCGCAGAGATGGAG 764
Db 261 ThrTrpAspProAlaGlnProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
QY 765 AAGGCCAAGGGGACCTCAGAGGAGCTCAATGATGAGAACTGGCATAGTGGTGGCT 824
Db 281 LysAlaLysGlyAsnProGluSerPheAsnAspGluAsnLeuArgIleValValAla 300
QY 825 GACCTGTTCTCTCGCGGATGGTGACCACTCGACCGCTGGCTGGCGGCTCTGCTC 884
Db 301 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 320
QY 885 ATGATCTTACATCCGATGTGAGCGCGCTGTCCAAACAGGAGATCGACGACCTGATAGG 944
Db 321 MetIleLeuHisProAspValGlnArgArgValGlnGlnIleAspAspValIleGly 340
QY 945 CAGGTGCGCGGACGAGAGTGGTGACCGAGGCTCAGTCCCTACACCACTCCCGTGATT 1004
Db 341 GlnValArgArgProGluMetGlyAspGlnAlaHisMetMetProTyrThrThrAlaValIle 360
```

```
QY 1005 CATGAGGTGCAGCGCTTTGGGGACATCGTCCCTGGGTGTGACCCATATGACATCCCGT 1064
Db 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 380
QY 1065 GACATCGAAGTACAGGGCTTCGCGATCCCTAAGGGAACGACACTCATCAACCACTGTCA 1124
Db 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400
QY 1125 TCGGTGCTGAAGGATGAGSCCGTCTGGGAGAAGCCCTCCGCTTCCACCCCGAACACTTC 1184
Db 401 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
QY 1185 CTGGATGCCCGAGGCGCACTTTGTGAAGCGGAGGCGCTTCTGCTTTTCTCAGCAGGCCGC 1244
Db 421 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 440
QY 1245 CGTGATGCTCGGGAGGCCCTGGCGCGCATGAGAGCTTCTCTTCTTCTCCTCCCTG 1304
Db 441 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrSerLeu 460
QY 1305 CTGCAGCACTTCAGCTTCTCGGTGCCCACTGGACAGCCCGCGCCAGCCACCATGGTGC 1364
Db 461 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 480
QY 1365 TTTGCTTTCTGCTGTCAGCCCATCCCGCTATGAGCTTTGTGCTGTGCCCCGC 1415
Db 481 PheAlaPheLeuValSerProSerProTyrGluLeuCysAlaValProArg 497
```

Search completed: February 25, 2004, 02:34:13
Job time : 88.6542 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 25, 2004, 02:12:38 ; Search time 12.7487 Seconds
(without alignments)
12555.283 Million cell updates/sec

Title: US-09-820-788A-1
Perfect score: 2886
Sequence: 1 cctgctggtctctgtgccc.....aaaaaaaaaaaaaaaaaaaaa 1537

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/cpn2.1/USPTO spool_p/US09820788/runat_24022004.141426.20026/app_query.fasta_1.12174
-DB=SwissProt 42 -QPMT=fastan -SUFFIX=rasp -MINMATCH=0.1 -LOPCU=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09820788 @CNC 1.1.143 @runat_24022004.141426.20026 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt 42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2282.5	79.1	497	CPD6 HUMAN	P10635 homo sapien
2	2140.5	74.2	497	CPDH_WACPA	Q29488 macaca fasc
3	2069.5	71.7	497	CPDJ_CALJA	O19992 callithrix
4	1736.5	60.2	499	CPDP_PIG	O4658 sus scrofa
5	1735	60.1	500	CPD4 RAT	P13108 rattus norv
6	1734.5	60.1	487	CPDE BOVIN	Q01361 bos taurus
7	1719	59.6	500	CPDI RAT	Q64680 rattus norv
8	1686.5	58.4	499	CPDF_CANFA	Q29473 canis famil
9	1653.5	57.3	500	CPDK_MESAU	Q9QY95 mesocricetu
10	1642.5	56.9	500	CPDR_MESAU	Q9QY96 mesocricetu
11	1642	56.9	500	CPD3 RAT	P12938 rattus norv
12	1640	56.8	500	CPD2 RAT	P10634 rattus norv
13	1619	56.1	504	CPD1 RAT	P10633 rattus norv
14	1616.5	56.0	504	CPD5 RAT	P12939 rattus norv
15	1611	55.8	500	CPDG_CAVPO	Q64403 cavia porce
16	1599.5	55.4	504	CPD9_MOUSE	P11714 mus musculu
17	1582	54.8	504	CPDA_MOUSE	P24456 mus musculu
18	1545.5	53.6	500	CPDS_MESAU	Q9QUJ1 mesocricetu

19	1457.5	50.5	505	1	CPDB_MOUSE	P24457 mus musculu
20	841	29.1	501	1	CPD6_MOUSE	O54750 mus musculu
21	837.5	29.0	502	1	CPJ2_HUMAN	P51589 homo sapien
22	828.5	28.7	502	1	CPJ3_RAT	P51590 rattus norv
23	818.5	28.4	491	1	CPB4_RABIT	P00178 oryctolagus
24	818.5	28.4	491	1	CPB5_RABIT	P12789 oryctolagus
25	813	28.2	501	1	CPJ5_MOUSE	O54749 mus musculu
26	811.5	28.1	500	1	CRJ1_RABIT	P52786 oryctolagus
27	808.5	28.0	491	1	CPB1_RAT	P00176 rattus norv
28	805.5	27.9	491	1	CPB9_MOUSE	P12790 mus musculu
29	801	27.8	494	1	CPCN_RAT	P24470 rattus norv
30	799.5	27.7	491	1	CPB2_RAT	P04167 rattus norv
31	797.5	27.6	490	1	CPB8_HUMAN	P10632 homo sapien
32	788.5	27.3	490	1	CPZ4_MOUSE	P56656 mus musculu
33	787.5	27.3	490	1	CPCK_MESAU	Q08078 mesocricetu
34	785.5	27.2	490	1	CPCK_MACFA	P33262 macaca fasc
35	785	27.2	490	1	CPCK_MESAU	P33263 mesocricetu
36	783	27.1	490	1	CPCT_MOUSE	Q64458 mus musculu
37	780	27.0	500	1	CPCB_RAT	P08683 rattus norv
38	779.5	27.0	493	1	CPB1_HUMAN	P05181 homo sapien
39	779.5	27.0	504	1	CPK1_ONCMY	Q92090 oncorhynch
40	779	27.0	491	1	CPF2_MOUSE	P33267 mus musculu
41	778	27.0	487	1	CPCL_CANFA	P56594 canis famil
42	778	27.0	492	1	CPBC_RAT	P33272 rattus norv
43	776.5	26.9	491	1	CPB6_HUMAN	P20813 homo sapien
44	775	26.9	490	1	CPCK_MESAU	P33264 mesocricetu
45	775	26.9	492	1	CPBJ_MOUSE	O55071 mus musculu

ALIGNMENTS

RESULT 1
CPD6_HUMAN
ID CPD6_HUMAN STANDARD; PRT; 497 AA.
AC P10635; Q16752;
DT 01-JUL-1989 (Rel. 11, Created)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome P450 2D6 (EC 1.14.14.1) (CYP2D6)
DE 4-hydroxylase).
GN CYP2D6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88314109; PubMed=3410476;
RA Gonzalez F.J., Vilbois F., Hardwick J.P., McBride O.W.,
Nebert D.W., Gelboin H.V., Meyer U.A.;
RT "Human debrisoquine 4-hydroxylase (P450IID1): cDNA and deduced amino acid sequence and assignment of the CYP2D locus to chromosome 22.";
RL Genomics 2:174-179(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=88122614; PubMed=3123397;
RA Gonzalez F.J., Skoda R.C., Kimura S., Umeno M., Zanger U.M.,
Nebert D.W., Gelboin H.V., Hardwick J.P., Meyer U.A.;
RT "Characterization of the common genetic defect in humans deficient in debrisoquine metabolism.";
RL Nature 331:442-446(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90072069; PubMed=2574001;
RA Kimura S., Umeno M., Skoda R.C., Meyer U.A., Gonzalez F.J.;
RT "The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene.";
RL Am. J. Hum. Genet. 45:889-904(1989).
RN [4]
RP VARIANT LYS-281 DEL (CYP2D6*9).
RX MEDLINE=93244880; PubMed=1844820;

RA Tyndale R., Aoyama T., Broly F., Matsunaga T., Inaba T., Kalow W.,
RA Gelboin H.V., Meyer U.A., Gonzalez F.J.;
RT "Identification of a new variant CYP2D6 allele lacking the codon
RT encoding Lys-281: possible association with the poor metabolizer
RL phenotype";
RL Pharmacogenetics 1:26-32(1991).
RN [5]
RP VARIANTS SER-34 AND THR-486 (CYP2D6*10).
RX MEDLINE=94115362; PubMed=8287064;
RA Yokota H., Tamura S., Furuya H., Kimura S., Watanabe M., Kanazawa I.,
RA Kondo I., Gonzalez F.J.;
RT "Evidence for a new variant CYP2D6 allele CYP2D6J in a Japanese
RT population associated with lower in vivo rates of sparteine
RT metabolism";
RL Pharmacogenetics 3:256-263(1993).
RN [6]
RP VARIANTS PRO-324 (CYP2D6*7).
RX MEDLINE=95147995; PubMed=7845481;
RA Evert B., Griese E.U., Eichelbaum M.;
RT "A missense mutation in exon 6 of the CYP2D6 gene leading to a
RT histidine 324 to proline exchange is associated with the poor
RT metabolizer phenotype of sparteine";
RL Naunyn Schmiedeberg Arch. Pharmacol. 350:434-439(1994).
RN [7]
RP VARIANT GLU-212 (CYP2D6*6B/6C).
RX MEDLINE=95172594; PubMed=7868129;
RA Daly A.K., Leathart J.B., London S.J., Idle J.R.;
RT "An inactive cytochrome P450 CYP2D6 allele containing a deletion and a
RT base substitution";
RL Hum. Genet. 95:337-341(1995).
RN [8]
RP VARIANT ILE-107 (CYP2D6*17).
RX MEDLINE=97126511; PubMed=8971426;
RA Masimirembwa C., Persson I., Bertilsson L., Hasler J.,
RA Ingelman-Sundberg M.;
RT "A novel mutant variant of the CYP2D6 gene (CYP2D6*17) common in a
RT black African population: association with diminished debrisoquine
RT hydroxylase activity";
RL Br. J. Clin. Pharmacol. 42:713-719(1996).
RN [9]
RP VARIANT ARG-42 (CYP2D6*12).
RX MEDLINE=96209916; PubMed=8655150;
RA Marex D., Legrand M., Sabbagh N., Lo-Guidice J.M., Boone P., Broly F.;
RT "An additional allelic variant of the CYP2D6 gene causing impaired
RT metabolism of sparteine";
RL Hum. Genet. 97:668-670(1996).
RN [10]
RP VARIANTS.
RX MEDLINE=97385645; PubMed=9241659;
RA Marex D., Legrand M., Sabbagh N., Guidice J.M., Spire C.,
RA Lafitte J.J., Meyer U.A., Broly F.;
RT "Polymorphism of the cytochrome P450 CYP2D6 gene in a European
RT population: characterization of 48 mutations and 53 alleles, their
RT frequencies and evolution";
RL Pharmacogenetics 7:193-202(1997).
RN [11]
RP VARIANT ARG-169 (CYP2D6*14).
RX MEDLINE=99164054; PubMed=10064570;
RA Wang S.L., Lai M.D., Huang J.D.;
RT "G169R mutation diminishes the metabolic activity of CYP2D6 in
RT Chinese";
RL Drug Metab. Dispos. 27:385-388(1999).
CC -!- FUNCTION: Responsible for the metabolism of many drugs and
CC environmental chemicals that it oxidizes. It is involved in the
CC metabolism of drugs such as antiarrhythmics, adrenoceptor
CC antagonists, and tricyclic antidepressants.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: By pregnancy.
CC -!- POLYMORPHISM: Highly polymorphic. Oxidative drug metabolism by
CC CYP2D6 is characterized by two phenotypes, the extensive
CC metabolizer (EM) and poor metabolizer (PM). Of the Caucasian

CC populations of Europe and North America, 5%-10% are of the PM
CC phenotype and are unable to metabolize the antihypertensive drug
CC debrisoquine and numerous other drugs.
CC -!- POLYMORPHISM: Allele CYP2D6*7 was also known as CYP2D6E, allele
CC CYP2D6*9 as CYP2D6C, allele CYP2D6*10 as CYP2D6J, allele CYP2D6*17
CC as CYP2D6Z.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -!- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;
CC NOTE=CYP2D6 alleles;
CC WWW="http://www.imm.ki.se/CYPalleles/cyp2d6.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M20403; AAA52153.1; -;
CC EMBL; X08006; CAA30807.1; -;
CC EMBL; M33388; AAA53500.1; -;
CC PIR; S01199; O4HDD1.
CC HSSP; P00179; 1DT6.
CC Genew; HGNC:2625; CYP2D6.
CC MIM; 124030; -;
CC InterPro; IPR001128; Cytochrome P450.
CC InterPro; IPR008069; EP450_CYP2D.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR01686; EP450ICYP2D.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Polymorphism.
FT METAL 443 443 IRON (HEME AXIAL LIGAND).
FT VARIANT 11 11 V -> M (in allele CYP2D6*35).
FT VARIANT 26 26 R -> H (in allele CYP2D6*21).
FT VARIANT 28 28 R -> C (in allele CYP2D6*22).
FT VARIANT 34 34 P -> S (in allele CYP2D6*10 and allele
FT CYP2D6*14; poor debrisoquine metabolism).
FT VARIANT 42 42 G -> R (in allele CYP2D6*12; impaired
FT metabolism of sparteine).
FT FTID=VAR_001256.
FT A -> V (in allele CYP2D6*23).
FT FTID=VAR_008369.
FT T -> I (in allele CYP2D6*17; poor
FT debrisoquine metabolism).
FT FTID=VAR_008337.
FT G -> R (in allele CYP2D6*14; poor
FT debrisoquine metabolism).
FT FTID=VAR_008338.
FT G -> E (in allele CYP2D6*6B and allele
FT CYP2D6*6C).
FT FTID=VAR_008339.
FT A -> S (in allele CYP2D6*33).
FT FTID=VAR_008370.
FT Missing (in allele CYP2D6*9).
FT FTID=VAR_008347.
FT R -> C (in allele CYP2D6*2, allele
FT CYP2D6*12, allele CYP2D6*14 and allele
FT CYP2D6*17; dbSNP:16947).
FT FTID=VAR_008340.
FT I -> L (in allele CYP2D6*24).
FT FTID=VAR_008371.
FT S -> L (in dbSNP:1800754).
FT FTID=VAR_014633.
FT H -> P (in allele CYP2D6*7; loss of
FT activity).
FT FTID=VAR_008348.

FT VARIANT 343 343 R -> G (in allele CYP2D6*25).
 FT /FTID=VAR 008372.
 FT VARIANT 369 369 I -> T (in allele CYP2D6*26).
 FT /FTID=VAR 008373.
 FT VARIANT 410 410 E -> K (in allele CYP2D6*27).
 FT /FTID=VAR 008374.
 FT VARIANT 486 486 S -> T (in allele CYP2D6*2, allele
 CYP2D6*10, allele CYP2D6*12, allele
 CYP2D6*14 and allele CYP2D6*17; impaired
 metabolism of sparteine).
 FT /FTID=VAR 008341.
 FT CONFLICT 374 374 M -> V (IN REF. 3).
 SQ SEQUENCE 497 AA; 55801 MW; 543F4D5F0E8DCAC CRC64;

Alignment Scores:
 Pred. No.: 7,37e-127 Length: 497
 Score: 2282.50 Matches: 443
 Percent Similarity: 89.54% Conservative: 2
 Best Local Similarity: 89.13% Mismatches: 1
 Query Match: 79.09% Indels: 51
 DB: 1 Gaps: 1

US-09-820-788A-1 (1-1537) x CPD6_HUMAN (1-497)

QY 78 ATGGGGCTAGAGCACTGGTGGCCCTGGCGGTAGTAGTGGCCATCTTCTGCTCTCTGGTG 137
 Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValAlaValAlaValAlaValAlaValAlaVal 20
 QY 138 GACCTGATCACCGCGCCCAACGCTGGCTGCACGCTACTCACCAGGCCCTGCCACTG 197
 Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTrpProGlyProLeuProLeu 40
 QY 198 CCGGGCTGGCAACCTGCTGATGGACTTCCAGAACACACCATCTGCTTCGACACAG 257
 Db 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
 QY 258 TTGGCGCGCGCTTCGGGACGTTTTCAGCTGACCTGGCTGCAGCGCGTGGTGGTGG 317
 Db 61 LeuArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
 QY 318 CTCATGGCTGGCGCGCTGGCGAGCGCTGTGACCCAGCGGAGGACACCCCGCCGAC 377
 Db 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
 QY 378 CGCCCGCTGTGCCATCACCGAGATCTGGGTTTGGCGCGCTGCCAA----- 428
 Db 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyValPhe 120
 QY 428 ----- 428
 Db 121 LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgArgPheSerValSerThrLeuArg 140
 QY 428 ----- 428
 Db 141 AsnLeuGlyLeuGlyLysSerLeuGluGlnTrpValThrGluGluAlaAlaCysLeu 160
 QY 429 -----GGACGCGCTTTCGCGCCCAACAGCTCTTCTGGACAAA 464
 Db 161 CysAlaAlaPheAlaAsnHisSerGlyArgProPheArgProAsnGlyLeuLeuAspLys 180
 QY 465 GCGTGAGCAACGTGATGCTCCCTCCATCCTGGGCGCGCTTCGAGTACGACGACCT 524
 Db 181 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTrpAspAspPro 200
 QY 525 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGAGTGAAGGAGGAGTCCGGCTTCTG 584
 Db 201 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluSerGlyPheLeu 220
 QY 585 CGGAGGCTGTAATGCTGCTCCCGTCTCCTGCTATATCCAGCGCTGGCTGGCAGGTC 644
 Db 221 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 240
 QY 645 CTACGCTTCCAAAGGCTTTCCTGACCCAGCTGGATGAGCTGCTAACTGAGCACGAGT 704

Db 241 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 260
 QY 705 ACCTGGGACCCAGCCAGCCCGGAGACCTGACTGAGGCGCTTCTGCGACAGATGGAG 764
 Db 261 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
 QY 765 AAGGCCAAGGGGAACCTCGAGACGAGCTTCAATGATGAGAACTCGCATAGTGGTGGCT 824
 Db 281 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 300
 QY 825 GACCTGTTCTCTGCGCGGATGGTGAACCTGACACCTGACACGCTGGCGCTGGGCGCTCTGCTC 884
 Db 301 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 320
 QY 885 ATGATCTTACATCCGATGTGCAGCGCGTGTCCAAACAGGAGATCGACGACCTGATAGG 944
 Db 321 MetIleLeuHisProAspValGlnArgArgValGlnGlnGlnGlnGlnGlnGlnGlnGln 340
 QY 945 CAGGTGCGCGGACACGAGATGGTGACGAGGCTCAATGCCCTACACCTGCGCTGATT 1004
 Db 341 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrAlaValIle 360
 QY 1005 CATGAGGTGCGAGCGCTTTGGGACATCGTCCCCCTGGGTGTGACCCATATGACATCCCGT 1064
 Db 361 HisGluValGlnArgPheGlyAspIleValProLeuGlyMetThrHisMetThrSerArg 380
 QY 1065 GACATCGAAGTACAGGCTTCGCGATCCTTAAGGAACGACACTCATCACCACCACTGTCA 1124
 Db 381 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 400
 QY 1125 TCGGTGCTCAAGGATGAGCGCTCTGGGAGAGCGCTTCCGCTTCCACCCCAACACTTC 1184
 Db 401 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 420
 QY 1185 CTGGATGCCAGGCGCCACTTTGTGAAGCGGAGGCGCTTCTGCTGCTTCTCAGCAGCGCCG 1244
 Db 421 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 440
 QY 1245 CGTGATCGCTCGGGGAGCCCTGCGCGCATGGAGCTTCTTCTTCTTCTTCTTCTTCTTCT 1304
 Db 441 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrSerLeu 460
 QY 1305 CTGACGACTTTCAGCTTCTCGGTGCCACTGACAGACCCCGCGCCAGCCACCATGGTGTG 1364
 Db 461 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 480
 QY 1365 TTGTGCTTCTGCTGACCCCATCCCTTATGAGCTTTGTGCTGCTGCTGCTGCTGCTGCT 1415
 Db 481 PheAlaPheLeuValSerProSerProTyrGluLeuCysAlaValProArg 497

RESULT 2
 ID CPDH MACFA STANDARD; PRT; 497 AA.
 AC Q29488;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 2D17 (EC 1.14.14.1) (CYP2D17).
 GN CYP2D17.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_taxid=95541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Lawton M.P., Laddison K.J., Speirs A.A., Mankowski D.C., Tweedie D.J.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By

CC similarity).
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38218; AAA79722.1; -
DR PIR; G02938; G02938.
DR HSP; P00179; 1D76.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 497 AA; 56010 MW; 3594AA88F04E58B1 CRC64;

Alignment Scores:

Pred. No.:	1.6e-118	Length:	497
Score:	2140.50	Matches:	417
Percent Similarity:	85.71%	Conservative:	9
Best Local Similarity:	83.90%	Mismatches:	20
Query Match:	74.17%	Indels:	51
DB:	1	Gaps:	1

US-09-820-788A-1 (1-1537) x CPD_MACFA (1-497)

QY	78	ATGGGGCTAGAGCACTGGTGGCCCTGGCTGATAGTGGCCATCTTCTGCTCTGCTG	137
DB	1	MetGluLeuAspAlaLeuValProLeuAlaValThrValAlaLeuLeuVal	20
QY	138	GACCTGATGACCGCGGCAACGCTGGCTGGCTGACCTACTACACAGGCCCTGCGCATG	197
DB	21	AspLeuMetHisArgGlnArgTrpAlaAlaArgTyrProProGlyProLeuLeu	40
QY	198	CCCGGGCTGGCAACCTGCTGATCTGGATCTCCAGACACACATCTGTTCCAGCAG	257
DB	41	ProGlyLeuGlyAsnLeuHisValAspPheLysAsnThrProTyrCysPheAspGln	60
QY	258	TTGCGGCGCGCTTCCGGGAGCTGTTACGCTGACCTGGCTGGCTGACCGCGCTGCTG	317
DB	61	LeuArgArgArgPheGlyAsnValPheSerLeuGlnLeuAlaTrpThrProValVal	80
QY	318	CTCAATGGCTGGCGCGCTGGCGAGCGCTGGTGGACCCAGCGGAGACACCGCCGAC	377
DB	81	LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrCysGlyGluAspThrAlaAsp	100
QY	378	CGCCGCTGTCGCCATCACCAGATCTGGTGGTGGTGGCGCGCTGCCAA-----	428
DB	101	ArgProProValProIleAsnGlnValLeuGlyPheGlyProArgSerGlnGlyValPhe	120
QY	428	-----	428
DB	121	LeuAlaArgTyrGlyProAlaTrpArgGluGlnArgArgPheSerValSerThrLeuArg	140
QY	428	-----	428
DB	141	AsnLeuGlyLeuGlyLysLysSerLeuGluGlnTrpValThrGluAlaAlaCysLeu	160
QY	429	-----CGAGCGCCCTTTCGCCCAACAGGCTCTTGGACACAA	464
DB	161	CysAlaAlaPheThrAspGlnAlaGlyArgProPheArgProAsnSerLeuLeuAspLys	180
QY	465	GCCGTGAGCAACGTGATCGCTTCCCTACCTCGCGGCGCGCTTCCAGTACGACACCT	524

DB	181	AlaValSerAsnValIleAlaSerLeuThrTyrGlyArgArgPheGluTyrAspAspPro	200
QY	525	CGCTTCTCAGGCTGGAGACTAGCTCAGAGGAGGAGTGAAGGAGTGGGCTTCTG	584
DB	201	ArgPheLeuArgLeuPheAspLeuThrHisGluAlaLeuLysGluGluSerGlyPheLeu	220
QY	585	CGCAGGTGCTGAATGCTCTCCCTGCTATATCCAGCGCTGGCTGGCAAGGTC	644
DB	221	ArgGluValLeuAsnAlaIleProLeuLeuLeuArgIleProGlyLeuAlaGlyLysVal	240
QY	645	CTACGCTTCCAAAAGGCTTCTCACCAGCTGGATGAGCTGCTAACTGACACAGGATG	704
DB	241	LeuArgSerGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet	260
QY	705	ACCTGGGACCCAGCCAGCCCCCGAGACCTGAGGCTTCTCTGGGAGAGATGGAG	764
DB	261	ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu	280
QY	765	AAGCCAAGGGGAACCTCAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGGCT	824
DB	281	LysAlaLysGlyAsnProGluSerSerPheAsnGluGluAsnLeuArgMetValValAla	300
QY	825	GACCTGTTCTCTCGCGGATGGTGACCACTCGACACGCTGGCTGGGCTCTCTGCTC	884
DB	301	AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTrpGlyLeuLeuLeu	320
QY	885	ATGATCTTACATCCGGATGTCAGCGCGCTGTCCAAAGAGAGATCGACGCTATAGG	944
DB	321	MetIleLeuHisProAspValGlnArgValGlnGlnGluIleAspAspValIleGly	340
QY	945	CAGTGGCGGACACAGAGATGGTGACCGCTCACATGCCCTACACACCTGCGCTGATT	1004
DB	341	GlnValArgArgProGluMetGlyAspGlnAlaArgMetProTyrThrAlaValIle	360
QY	1005	CATGAGTGGACGCTTTGGGACATCGTCCCCCTGGGTGGACCATATGACATCCCGT	1064
DB	361	HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg	380
QY	1065	GACATCGAAGTACAGGGCTTCCGATCCCTAAGGAGACGACCTCATCACCACTGTCA	1124
DB	381	AspIleGluLeuGlnGlyPheLeuIleProLysGlyThrThrLeuPheThrAsnLeuSer	400
QY	1125	TGCGTGTGAAGGATGAGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCGAAACATTC	1184
DB	401	SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe	420
QY	1185	CTGATGCCAGGCGCACTTTGTGAAGCGGAGCGCTTCTGCTTTCTACGAGGCGCG	1244
DB	421	LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg	440
QY	1245	CGTCATGCTCGGGGAGCCCTGGCCGGATGGAGCTTCTTCTTCTTCCCTCCCTG	1304
DB	441	ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrCysLeu	460
QY	1305	CTGACGACCTTACGCTTCTCGGTGCCACTCGACAGCCCGCGCCAGCCACCATGGTCTC	1364
DB	461	LeuGlnArgPheSerPheSerValProAlaGlyGlnProArgProSerHisGlyVal	480
QY	1365	TTTCTTCTTCTGGTACCCCATCCCTATGAGCTTTGTGTGTGTGTGTGTGTGTGTGT	1415
DB	481	PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg	497
RESULT 3			
CPD	CALJA	STANDARD;	PRT; 497 AA.
ID	CPDJ	CALJA	AC
AC	O18992;		
DT	15-DEC-1998	(Rel. 37, Created)	
DT	15-DEC-1998	(Rel. 37, Last sequence update)	
DT	28-FEB-2003	(Rel. 41, Last annotation update)	
DE	Cytochrome P450 2D19	(EC 1.14.14.1) (CYF1D19) (P450 CM2D-1).	
GN	CYP2D19.		
OS	Callithrix jacchus (Common marmoset).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
 OC Callitrix.
 OX NCBI_TaxID=9483;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Liver;
 RC MEDLINE=97223367; PubMed=9056237;
 RA Igarashi T., Sakuma T., Isogai M., Nagata R., Kamataki T.;
 RT "Marmoset liver cytochrome P450s: study for expression and molecular
 cloning of their cDNAs.";
 RL Arch. Biochem. Biophys. 339:85-91(1997).
 CC -1- FUNCTION: RESPONSIBLE FOR THE METABOLISM OF MANY DRUGS AND
 ENVIRONMENTAL CHEMICALS THAT IT OXIDIZES.
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 oxidized flavoprotein + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -1- INDUCTION: P450 can be induced to high levels in liver and other
 tissues by various foreign compounds, including drugs, pesticides,
 and carcinogens.
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; D29822; BAA22155.1; -;
 DR HSSP; P00179; 1D76.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR008069; EP450_CYP2D.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01686; EP450ICYP2D.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 497 AA; 55911 MW; A482AE71EAD6CAF CRC64;
 Alignment Scores:
 Pred. No.: 2,35e-114 Length: 497
 Score: 2059.50 Matches: 405
 Percent Similarity: 83.70% Conservative: 11
 Best Local Similarity: 81.49% Mismatches: 30
 Query Match: 71.71% Indels: 51
 DB: 1 Gaps: 1
 US-09-820-788a-1 (1-1537) x CPDJ_CALJA (1-497)
 QY 78 ATGGGGCTAGACACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCTGCTCTGGTG 137
 Db 1 MetGlyLeuAspAlaLeuValProLeuAlaValThrValAlaIlePheValLeuVal 20
 QY 138 GACCTGATGACCGCGCCCAACGCTGGCTGCACTGCTACTCACCAGGCCCTGGCCACTG 197
 Db 21 AspLeuMetHisArgArgGlnArgTTPAlaAlaArgTyProProGlyProMetProLeu 40
 QY 198 CCGGGCTGGCAACTGCTGCAATGGACTTCCAGAACACACACAPACTGCTTCGACACAG 257
 Db 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProAsnSerPheAsnGln 60
 QY 258 TTGGCGCGCGCTTGGGAGCGTGTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 317
 Db 61 LeuArgArgPheGlyAspValPheSerLeuGlnLeuAlaIlePheProValValVal 80
 QY 318 CTCATATGGGTGGCGCGCTGGCGGCGCTGGTGACCCACCGCGGAGGACACCGCCGAC 377
 Db 81 LeuAsnGlyLeuGluAlaValArgGluAlaLeuValThrArgGlyGluAspThrAlaAsp 100
 QY 378 CGCCCGCTGTGCCCATCACCCAGATCTCTGGGTTTGGGCGCGGTGCCCAA----- 428

Db 101 ArgProValProIleThrGluMetLeuGlyPheGlyProHisSerGlnGlyLeuPhe 120
 QY 428 ----- 428
 Db 121 LeuAlaArgTyGlyProAlaTrpArgGluGlnArgArgPheSerValSerThrLeuArg 140
 QY 428 ----- 428
 Db 141 AsnLeuGlyLeuGlyLysLysSerLeuGluGlnTrpValThrGluGluAlaThrTyLeu 160
 QY 429 -----GGAGCCCTTTTCGCCCCCAACGGTCTCTTGACAAA 464
 Db 161 CysAlaAlaPheAlaAspHisAlaGlyArgProPheArgProAsnGlyLeuLeuAspLys 180
 QY 465 GCGGTGAGCAACGTCATCGCTCCCTCCTCCTCGGCGCGCTTCGAGTACGACGACCTT 524
 Db 181 AlaValSerAsnValIleAlaSerLeuThrCysArgArgArgPheGluTyTrpAsnAspPro 200
 QY 525 CGCTTCTCAGGCTGCTGGACCTAGCTCAGAGGAGCTGAAGGAGAGTGGGCTTCTG 584
 Db 201 CysLeuLeuArgLeuLeuAspLeuThrMetGluGlyLeuLysGluGluSerGlyLeuLeu 220
 QY 585 CCGAGGTGCTCAATGCTGTCCCTCCTCTGTCATATCCAGAGCGCTGGCTGGCAAGTTC 644
 Db 221 ArgGluValLeuAsnAlaIleProValLeuLeuArgIleProGlyLeuAlaGlyLysVal 240
 QY 645 CTACGCTTCCAAAGGCTTCTCCTGACCCAGCTGGATGAGTCTTAACCTGAGCAGCAGGATG 704
 Db 241 LeuArgSerGlnLysAlaPheLeuAlaGlnLeuAspGluLeuLeuThrGluHisArgMet 260
 QY 705 ACCTGGGACCCAGCCCGAGAGCTGAGTCTGAGGCTTCTTCTGGCAGAGATGGAG 764
 Db 261 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 280
 QY 765 AAGGCCAAGGGAAACCTGAGAGAGCTTCAATGATGAGAACTGGCGCATGTGGTGGCT 824
 Db 281 LysThrLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuHisLeuValValAla 300
 QY 825 GACCTGTTCTCTGCGGGATGTGACCTGACCTGACCTGCGCTGGCGCTTCTTCTGCTC 884
 Db 301 AspLeuPheSerAlaGlyMetValThrThrSerIleThrLeuAlaTrpGlyLeuLeuLeu 320
 QY 885 ATGATCTACATCCGATGTGCGCGCTGTCCAAACAGGAGATCGACAGCTGTATAGG 944
 Db 321 MetIleLeuHisProAspValGlnArgValGlnGlnGluLeuAspValIleGly 340
 QY 945 CAGGTGGCGGACACAGAGATGGGTGACCGAGGTCAATGCCCTACACCACTCCGCTGATT 1004
 Db 341 ArgValArgArgProGluMetGlyAspGlnThrTyMetProTyThrThrAlaValIle 360
 QY 1005 CATGAGGTGACGCGCTTTGGGACATCGTCCCTGGGTGTGACCATATGACATCCCTG 1064
 Db 361 HisGluValGlnArgPheAlaAspIleValProLeuGlyValThrHisMetThrSerArg 380
 QY 1065 GACATCGAAGTACAGGCTTCCGATCCCTAAGGAAGACACTCATCACCAACCTGTCA 1124
 Db 381 AspIleGluValGlnGlyPheLeuIleProLysGlyThrThrLeuPheThrAsnLeuSer 400
 QY 1125 TCGGTGTGAAGGATGAGCGCTTGGGAGAACGCCCTTCCGCTTCCACCCCGAACACTTC 1184
 Db 401 SerValLeuLysAspGluAlaAsnTrpGluLysProPheArgPheHisProGluHisPhe 420
 QY 1185 CTGGATGCCAGGGCCACTTTGTGAAGCGGAGGCTTCTCGCTTCTTCTCAGCAGGCCG 1244
 Db 421 LeuAspAlaGlnGlyArgPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 440
 QY 1245 CGTCATGCTCGGGAGGCCCTGCGCCGCGCATGGAGCTTCTCTTCTTCTACCTCCCTG 1304
 Db 441 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrCysLeu 460
 QY 1305 CTGACGACTTCAAGTCTTCTCGGTGCCACTGAGACAGCCCGGCCCGCAGCCCATGGTGT 1364


```
Db 115 PheGlyProArgSerGlnGlyValValLeuAlaAArgTyrGlyProAlaTrpArgGlnGln 134
QY 428 ----- 428
Db 135 ArgArgPheSerValSerThrPheArgHisPheGlyLeuGlyLysSerLeuGluGln 154
QY 429 -----GGAGCGCCC 437
Db 155 TrpValThrGluAlaArgCysLeuCysAlaAlaPheAlaAAspHisSerGlyPhePro 174
QY 438 TTTCGCCCAACGGTCTCTTGACAAACCGTGAACAGCTGATCGCTCCCTCCACCTGC 497
Db 175 PheSerProAsnThrLeuLeuAspLysAlaValCysAsnValIleAlaSerLeuLeuPhe 194
QY 498 GGGCGCCCTCCAGTAGACAGACCTCGCTTCCTCAGCGCTGCTGGACTAGCTCAGAG 557
Db 195 AlaCysArgPheGluTyrAsnAspProArgPheIleArgLeuLeuAspLeuLysAsp 214
QY 558 GGACTGAAGGAGGAGTCCGGCTTCTCGCGAGGCTGCTGAATGCTGCTCCCGTCTCCTG 617
Db 215 ThrLeuGluGluGluSerGlyPheLeuProMetLeuLeuAsnValPheProMetLeuLeu 234
QY 618 CATATCCAGCGCTGGTGGCAAGTCTCTACGCTTCCAAAGGCTTCTCCTGACCCAGCTG 677
Db 235 HistileProGlyLeuLeuGlyLysValPheSerGlyLysLysAlaPheValAlaMetLeu 254
QY 678 GATGAGCTGCTAATCAGACAGAGTACCTGGGACCCAGCCAGCCCGCCGAGACCTG 737
Db 255 AspGluLeuLeuThrGluHisLysValThrTrpAspProAlaGlnProProArgAspLeu 274
QY 738 ACTGAGGCTTCTCGGACAGATGAGAGAGCCAGGGGACCCCTGAGAGCAGCTTCAAT 797
Db 275 ThrAspAlaPheLeuAlaGluValGluLysAlaLysGlyAsnProGluSerPheAsn 294
QY 798 GATGAGAACCTCGCATAGTGTGCTGACCTGTTCTGCGCGGATGCTGACACCTCG 857
Db 295 AspGluAsnLeuArgValValAlaAlaAspLeuPheMetAlaGlyMetValThrThrSer 314
QY 858 ACCAGCTGGCTGGGGCTCTGCTCATGATCTCATGATCTCATGATCTGATGATGATGATG 917
Db 315 ThrThrLeuThrTrpAlaLeuLeuPheMetIleLeuHisProAspValGlnCysArgVal 334
QY 918 CAACAGGAGATCGACGATGATGAGGAGGAGTGGCGGACAGAGATGGGTGACAGGCT 977
Db 335 GlnGlnGluIleAspGluValIleGlyGlnValArgProGluMetAlaAspGlnAla 354
QY 978 CACATGCCCTACACCACTGCGCTGATTCATGATGATGATGATGATGATGATGATGATG 1037
Db 355 ArgMetProPheThrAsnAlaValIleHisGluValGlnA-gPheAlaAspIleLeuPro 374
QY 1038 TTGGGTGTGACCCATATGACATCCCGTGCATCGAATCGAAGTACGAGGCTTCGCGATCCCTAAG 1097
Db 375 LeuGlyValProHisLysThrSerArgAspIleGluValGlnGlyPheLeuIleProLys 394
QY 1098 GGAACGACACTCATCACCACCTGTCATGCTGCTGAGAGTATGAGCGCTGGGAGAG 1157
Db 395 GlyThrThrLeuIleThrAsnLeuSerSerValLeuLysAspGluThrValTrpGluLys 414
QY 1158 CCCTTCCGCTTCCACCCGCAACACTTCTGATGATGATGATGATGATGATGATGATGATG 1217
Db 415 ProLeuArgPheHisProGluHisPheLeuAspAlaGlnGlyAsnPheValIysHisGlu 434
QY 1218 GCCTTCCTGCTTCTTACAGAGCGCGCTGCATGCTCGGAGAGCCCTGCGCGCATG 1277
Db 435 AlaPheMetProPheSerAlaGlyArgAlaCysLeuGlyGluProLeuAlaArgMet 454
QY 1278 GAGCTTCTTCTTCTTCACTCCCTGCTGAGACACTTCACTTCTGCTGCTGCTGCTGCTGCTG 1337
Db 455 GluLeuPheLeuPheThrCysLeuLeuGlnArgPhePheSerValProThrGly 474
QY 1338 CAGCCCCCGCCGAGCAGCAGCATGGTCTTTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTG 1397
```

```
Db 475 GlnProArgProSerAspTyrGlyIlePheGlyAlaLeuThrThrProArgProTyrGln 494
QY 1398 CTTTGTGCTGCGCCCGC 1415
Db 495 LeuCysAlaSerProArg 500
RESULT 6
CPDE_BOVIN
ID_CPDE_BOVIN STANDARD; PRT; 487 AA.
AC Q01361;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D14 (EC 1.14.14.1) (CYPIID14) (Fragment).
GN CYP2D14.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93011103; PubMed=1396578;
RA Tsuneoka Y., Matsuo Y., Higuchi R., Ichikawa Y.;
RT "Characterization of the cytochrome P-450IID subfamily in bovine
RL liver. Nucleotide sequences and microheterogeneity.";
CC Eur. J. Biochem. 208:739-746(1992).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X68013; CAA48149.1; -
CC HSP; P00179; I076.
CC InterPro; IPR001128; Cytochrome_P450.
CC InterPro; IPR008069; EP450_CYP2D.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR01686; EP450ICYP2D.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT NON TER 1
FT METAL 433 433 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 487 AA; 54687 MW; D61CFE3BBADE19E7 CRC64;
Alignment Scores:
Pred. NO.: 1.09e-94 Length: 487
Score: 1734.50 Matches: 338
Percent Similarity: 77.62% Conservative: 40
Best Local Similarity: 69.40% Mismatches: 58
Query Match: 60.10% Indels: 51
DB: 1
US-09-820-788A-1 (1-1537) x CPDE_BOVIN (1-487)
```


RC STRAIN=Beagle;
RX MEDLINE=98162950; PubMed=95044424;
RA Tasaki T., Nakamura A., Itoh S., Ohashi K., Yamamoto Y., Masuda M.,
RA Iwata H., Kazusaka A., Kamataki T., Fujita S.;
RT "Expression and characterization of dog CYP2D15 using baculovirus
RT expression system";
RJ J. Biochem. 123:162-168(1998).
RN [3]
RN CHARACTERIZATION.
RP MEDLINE=98389575; PubMed=9721180;
RA Roussel F., Duignan D.B., Lawton M.P., Obach R.S., Strick C.A.,
RA Tweedie D.J.;
RT "Expression and characterization of canine cytochrome P450 2D15.";
RT Arch. Biochem. Biophys. 357:27-36(1998).
CC -1- FUNCTION: HIGH ACTIVITY FOR THE HYDROXYLATION OF BUNITROL AND
CC IMIPRAMINE; LOW ACTIVITY ON DEBRISOQUINE.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2O).
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- TISSUE SPECIFICITY: LIVER. ALSO DETECTED IN SEVERAL OTHER TISSUES.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D17397; BAA04220.1; -;
DR EMBL; AB004268; BAA20357.1; -;
DR F1R; JC4157; J04157.
DR HSSP; P00179; 1D76.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT INIT_MET 0
FT METAL 445 445 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 499 AA; 56301 MW; 27E352B5B309E7F1 CRC64;

Alignment Scores:
Pred. No.: 7.18e-92 Length: 499
Score: 1686.50 Matches: 331
Percent Similarity: 75.15% Conservative: 41
Best Local Similarity: 66.87% Mismatches: 70
Query Match: 58.44% Indels: 53
DB: 1 Gaps: 3

US-09-820-788A-1 (1-1537) x CPDF_CANFA (1-499)
QY 87 GAAGCACTGGTCCCTGGCGGTGAGTGGCCATCTTCTGCTCTGCTGGTGGACCTGATG 146
DB 6 AspThrLeuGlyProLeuAlaValAlaValAlaLeuPheLeuLeuValAspLeuMet 25
QY 147 CACCGCGCCAAAGCTGGCTGCACGCTACTACACAGGCCCTCCACTGCGCGGGTGG 206
DB 26 HisArgArgArgArgTrpAlaThrArgTyrProProGlyProThrProValProMetVal 45
QY 207 GCGACCTCTGATGTGACTTCAGACTTCAGACACACCATATCTGC---TTGACACAGTTGGG 263
DB 46 GlyAsnLeuLeuGlnMetAspPheGlnGlu---ProIleCysTyrPheSerGlnLeuGln 64
QY 264 CGCGCGCTTCGGGAGCTGTTTACGCTGACGCTGGCTGGCGACCGCTGGTGTGCTCAAT 323
DB 65 GlyArgPheGlyAsnValPheSerLeuGluLeuAlaTrpThrProValValValLeuAsn 84
QY 324 GGGCTGGCGCGCTGGCGAGGCGTGGTGACCCACCGCGAGGACACCGCGCGCGCG 383

Db 85 GlyLeuGluAlaValArgGluAlaLeuValHisArgSerGluAspThrAlaAspArgPro 104
QY 384 CTGTGGCCCATCACAGATCTGGTGGTGGTGGCGCGTTCCTCAAA----- 428
Db 105 ProMetProIleTyrAspHisLeuGlyLeuGlyProGluSerGlnGlyLeuPheLeuAla 124
QY 428 ----- 428
Db 125 ArgTyrGlyArgAlaTrpArgGluGlnArgArgPheSerLeuSerThrLeuArgAsnPhe 144
QY 428 ----- 428
Db 145 GlyLeuGlyArgLysSerLeuGluGlnTrpValThrGluGluAlaSerCysLeuCysAla 164
QY 429 -----GGAGCGCCCTTTCGCCCAACGGTCTCTTGGACAAAGCCGTG 470
Db 165 AlaPheAlaGluGlnAlaGlyArgProPheGlyProGlyAlaLeuLeuAsnLysAlaVal 184
QY 471 ACACACGTGATCCCTCCCTCACCTCGCGGCGCGCTTCGAGTACGACGACCTCGCTTC 530
Db 185 SerAsnValIleSerSerLeuThrTyrGlyArgArgPheGluTyrAspAspProArgLeu 204
QY 531 CTCAGCTGTGGACCTAGCTCAGAGGAGCTGAAGAGAGAGCTGGGCTTCTGCGCGAG 590
Db 205 LeuGlnLeuLeuGluLeuThrGlnGlnAlaLeuLysGlnAspSerGlyPheLeuArgGlu 224
QY 591 GTGCTGAATGTCTCCCGTCTCTGATATCCAGCGCTGGTGGCGAAGGTCTCTACGC 650
Db 225 AlaLeuAsnSerIleProValLeuLeuHisIleProGlyLeuAlaSerLysValPheSer 244
QY 651 TTCCAAAAGGCTTCTCTGACCCAGCTGGATGAGCTGCTAACTGAGCAGAGATGACCTGG 710
Db 245 AlaGlnLysAlaIleIleThrLeuThrAsnGluMetIleGlnGluHisArgLysThrArg 264
QY 711 GACCAGCCCGAGCCCGGAGACCTGACTGAGGCTTCTCTGCGAGAGATGAGAGAGGCC 770
Db 265 AspProThrGlnProProArgHisLeuIleAspAlaPheValAspGluIleGlyLysAla 284
QY 771 AAGGGAAACCTGAGAGAGCTTCAATGATGAGAACCTCGGCATAGTGTGCTGACCTCG 830
Db 285 LysGlyAsnProLysThrSerPheAsnGluGluAsnLeuLysMetValThrSerAspLeu 304
QY 831 TTCTCTGCGGGGATGGTGACCACTCGACCGCTGGGCTGGGCTCTCTGCTCATGATC 890
Db 305 PheIleAlaGlyMetValSerThrSerIleThrLeuThrTrpAlaLeuLeuMetIle 324
QY 891 CTACATCCGATGTGTCAGCGCGTGTCCAAACAGAGATCGACGAGTATAGGCGAGTGG 950
Db 325 LeuHisProAspValGlnArgArgValGlnGlnGluIleAspGluValIleGlyArgGlu 344
QY 951 CGCGACCCAGAGATGGGTGACCGCTCACATGCCCTACACCTGCGCGTGGTTCATGAG 1010
Db 345 GlnLeuProGluMetGlyAspGlnThrArgMetProPheThrValAlaValIleHisGlu 364
QY 1011 GTGACGCTTGGGGACATGTCCTCCCTGGGTGTGACCCATATGACATCCCGTGCACATC 1070
Db 365 ValGlnArgPheGlyAspIleValProLeuGlyValProHisMetThrSerArgAspThr 384
QY 1071 GAAGTACAGGGCTTCCGATCCCTTAAGGAAACGACATCATCACCACTGTCATCGGTG 1130
Db 385 GluValGlnGlyPheLeuIleProLysGlyThrThrLeuIleThrAsnLeuSerSerVal 404
QY 1131 CTGAGGATGAGCGCTCTGGGAGNAGCCCTCCGCTCCACCCCGACACTTCTCTGGAT 1190
Db 405 LeuLysAspGluLysValTrpLysLysProPheArgPheTyrProGluHisPheLeuAsp 424
QY 1191 GCCCAGGCGCACTTTGTGAAGCGGAGGCTTCTCTGCTCTTCTCAGCAGGCGCGCTGCA 1250
Db 425 AlaGlnGlyHisPheValLysHisGluAlaPheMetProPheSerAlaGlyArgVal 444
QY 1251 TGCTTCGGGAGCGCTCTGGCGCGCATGAGCTTCTCTCTTCTTCTTCTCTCTCTCTCT 1310

Db 445 CysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrCysLeuLeuGln 464
 QY 1311 CACTTCAGTTCTCGGTGCCCACTGGACAGCCCGCGCCAGCCACATGGTGTCTTTGCT 1370
 Db 465 ArgPheSerPheSerValProAlaGlyGlnProArgProSerAspHisGlyValPheThr 484
 QY 1371 TTCTGGTGTGACCCCATCCCTATCAGCTTTGTGTGTGTCGCCCGC 1415
 Db 485 PheLeuLysValProAlaProPheGlnLeuCysValGluProArg 499

RESULT 9

ID_CPDK_MESAU STANDARD; PRT; 500 AA.
 AC Q90YG5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 2D20 (EC 1.14.14.-) (CYP11D20).
 GN CYP2D20.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_TaxID=10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=20533996; PubMed=11083025;
 RA Oka T., Fukuhara M., Ushio F., Kurose K.;
 RT "Molecular cloning and characterization of three novel cytochrome
 P450 2D isoforms, CYP2D20, CYP2D27, and CYP2D28 in the Syrian hamster
 (Mesocricetus auratus).";
 RT (Mesocricetus auratus).";
 RL Comp. Biochem. Physiol. 127C:143-152(2000).
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AB031864; BAA89313.1; -;
 DR HSP; P00179; 1D76.
 DR InterPro; IPR001128; Cytochrome P450.
 DR InterPro; IPR008069; EP450_CYP2D.
 DR Pfam; PF00067; p450; 1.
 DR PRINTS; PR01686; EP450ICYP2D.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Mitochondrion; Endoplasmic reticulum.
 FT METAL 446 446 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 500 AA; 56503 MW; 9948102706C50144 CRC64;

Alignment Scores:

Pred. No.: 6,228-90 Length: 500
 Score: 1653.50 Matches: 324
 Percent Similarity: 73.31% Conservative: 33
 Best Local Similarity: 66.53% Mismatches: 79
 Query Match: 57.29% Indels: 51
 DB: 1 Gaps: 1

US-09-820-788a-1 (1-1537) x CPDK_MESAU (1-500)

QY 108 GTGATAGTGGCCATCTTCTGTCTGTGGACCTGATGACCGCGCAACGCTGGGCT 167
 Db 14 IlePheThrAlaLeuPheLeuLeuValAspLeuMetHisArgLysPheTrpArg 33

QY 168 GCACGCTACTCACAGGCCCTCTCCACTGCCCCGGGCTGGCAACCTGCTCATGTGAC 227
 Db 34 AlaArgTyrProGlyProMetLeuProGlyLeuGlyAsnLeuLeuValAsp 53
 QY 228 TTCAGAACACACATCTGCTTCGACAGTTGGCGCGCGCTTCGGGAGCGTGTTCAGC 287
 Db 54 PheGluAsnMetProTyrSerLeuTyrLysPheGlnArgTyrGlyAspValPheSer 73
 QY 288 CTGAGCTGGCTCGACCGCGGTGCTGCTCAATGGCTGGCGCGCGTGGCGAGCGC 347
 Db 74 LeuGlnMetAlaTrpLysProValValIleAsnGlyLeuLysAlaValArgGluVal 93
 QY 348 CTGCTGACCCACGCGGAGGACACCGCGCGCGCTGTGCCCATCACCCAGATCTG 407
 Db 94 LeuValAsnCysGlyGluAspThrAlaAspArgProValProIlePheAsnHisLeu 113
 QY 408 GGTTTTGGCGCGCTTCCCAA----- 428
 Db 114 GlyTyrArgProLysSerGlnGlyValValPheAlaArgTyrGlyProGlnTrpArgGlu 133
 QY 428 ----- 428
 Db 134 GlnArgArgPheSerValSerThrMetArgAspPheGlyValGlyLysLysSerLeuGlu 153
 QY 429 -----GGAGCG 434
 Db 154 GlnTrpValThrGluGluAlaGlyHisLeuCysAspAlaPheThrGlnGluAlaGlyHis 173
 QY 435 CCTTTGCGCCCAACGGTCTCTCGACAAAGCCGCTGAGCAACGTGATCGCTCCTCACC 494
 Db 174 ProPheAsnProIleThrLeuLeuAsnLysSerValCysAsnValIleSerSerLeuIle 193
 QY 495 TCGCGGCGCGCTTCGATGACGACCGCTGCTTCTCAGGCTGTGGAGCTAGCTCAG 554
 Db 194 TyrAlaHisArgPheAspTyrGluAspProPhePheAsnLysLeuLeuLysThrLeuGln 213
 QY 555 GAGGACTGAAGGAGGAGTGGGCTTTCTGGCGGAGGTGCTGAATGCTGCTCCCTCCTC 614
 Db 214 GluSerPheGlyGluAspSerGlyPheIleAlaGluValLeuAsnAlaValProValLeu 233
 QY 615 CTGCATATCCAGCGCTGGCTGGCAAGCTCTACGCTTCCAAAGGCTTCTCAGCCAC 674
 Db 234 LeuArgIleProGlyLeuProGlyLysAlaPheProLysLeuThrAlaPheMetAspSer 253
 QY 675 CTGATGAGCTGTAATGACGACAGGATGACCTGGGACCCAGCCCGCCCGGAGAC 734
 Db 254 LeuTyrLysMetLeuIleGluHisLysThrThrTrpAspProAlaGlnProProArgGly 273
 QY 735 CTGACTGAGGCTTCTGGCAGAGATGGAGAAGCCCAAGGGGACCTTGAGAGAGCTTC 794
 Db 274 LeuThrAspAlaPheLeuAlaGluValIleGlyLysAlaLysGlyArgProGluSerPhe 293
 QY 795 AATGATGAGAACCTGCGCATAGTGGCTGACCTGTTCTCTGCGGGATGGTGACAC 854
 Db 294 AsnAspGluAsnLeuHisValValValAlaAspLeuPheIleAlaGlyMetValThr 313
 QY 855 TCGACCACTGCGCTGGGCGCTCCTGCTCATGATCTTACATCCGAGTGTGACGCGCT 914
 Db 314 SerThrThrLeuSerTrpAlaLeuLeuMetIleLeuHisProAspValGlnSerArg 333
 QY 915 GTCCACAGGAGATCGAGCGCTATAGGCGAGGTGGCGGACCCAGAGATGGTGAC 974
 Db 334 ValGlnGlnGluIleAspAspValIleGlyGlnValArgArgProGluMetAlaAspGln 353
 QY 975 GCTCACATGCCCTACACCACTCCGCTGATTCATCAGGTGACGCGCTTTGGGACATGTC 1034
 Db 354 AlaArgMetProTyrThrAsnAlaValIleHisGluValGlnArgPheGlyAspIleAla 373
 QY 1035 CCCCTGGGTGACCCATATGACATCCCGTACATCCGAAGTACAGGCTTCGCGATCCCT 1094
 Db 374 ProValAsnValProHisMetThrSerArgAspValGlnValGlnGlyPheLeuIlePro 393

QY 1095 AAGGGAACGACACTATACCAACCTGTCTATCGGTGCTGAAGATGAGCGCTCTCGGAG 1154
Db 394 LygGlyThrThrLeuLeuProAsnLeuSerSerValLeuLysAspGluThrValTTPGlu 413
QY 1155 AAGCCCTTCGCTTCCACCCCGGACACTTCTCGATGCGCCAGGCGCACTTGTGAGCCG 1214
Db 414 LyProLeuHisPheHisProGluHisPheLeuAspAlaGlnGlyArgPheValLysGln 433
QY 1215 GAGGCGCTTCTGCTTCTCAGCAGGCGCGCTGTCATGCTCGGGGAGCCCTTGGCGCCG 1274
Db 434 GluAlaPheMetProPheSerAlaGlyArgAlaCysLeuGlyGluProLeuAlaArg 453
QY 1275 ATGAGCTTCTCTTCTTCTTCTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1334
Db 454 MetGluLeuPheLeuPheThrCysLeuLeuGlnArgPheSerPheSerValProAla 473
QY 1335 GGACAGCCCGCGCCAGCCACCATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1394
Db 474 GlyGlnProArgProSerAspGlnGlyValPheAlaLeuProValThrProThrProTyr 493
QY 1395 GAGCTTTGTGCTGCTGCGCCG 1415
Db 494 GluLeuCysAlaValValArg 500

RESULT 10
CPDR_MESAU
ID CPDR_MESAU STANDARD; PRT; 500 AA.
AC O90Y6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D27 (EC 1.14.14.-) (CYP11D27).
GN CYP2D27.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=2053396; PubMed=11083025;
RA Oka T., Fukuhara M., Ushio F., Kurose K.;
RT "Molecular cloning and characterization of three novel cytochrome
P450 2D isoforms, CYP2D20, CYP2D27, and CYP2D28 in the Syrian hamster
(Mesocricetus auratus).";
RL Comp. Biochem. Physiol. 127C:143-152(2000).
CC -!- FUNCTION: Has bufuralol 1'-hydroxylase and debrisoquine 4-
hydroxylase activities.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By
similarity).
CC -!- TISSUE SPECIFICITY: Expressed in liver, but not in kidney, small
intestine, and brain.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR ENBL; AB031863; BAA89312.1; -.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008069; BP450_CYP2D.
DR Pfam; PF00067; P450_1.
DR PRINTS; PR01686; BP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;

KW Microsome; Endoplasmic reticulum.
FT METAL 446 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 500 AA; 56489 MW; 7A5DE5878F97A954 CRC64;

Alignment Scores:
Pred. No.: 2,75e-89 Length: 500
Score: 1642.50 Matches: 320
Percent Similarity: 73.10% Conservative: 36
Best Local Similarity: 65.71% Mismatches: 80
Query Match: 56.91% Indels: 51
DB: 1 Gaps: 1

US-09-820-788A-1 (1-1537) x CPDR_MESAU (1-500)

QY 108 GTGATAGTGGCCATCTTCTGCTCTGCTGGACCTGATGCACCGCGCCACACGCTGGGCT 167
Db 14 IlePheThrAlaLeuPheLeuLeuValAspLeuMetHisArgArgLysPheTrpArg 33
QY 168 GCACCTACTACACAGGCCCCCTGCATGCTGCCCGGCTGGGCAACCTGCTGATGTGAC 227
Db 34 AlaArgTyrProProGlyProMetProLeuProGlyLeuGlyAsnLeuLeuValAsp 53
QY 228 TTCAGAACACACACACTACTGCTTCCACAGTGTGCGCGCGCTTCCGGGACGTTTCAGC 287
Db 54 PheGluHisMetProTyrSerLeuTyrLysPheArgGlnArgTyrGlyAspValPheSer 73
QY 288 CTGCAGCTGGCTGACCGCGCTGCTGCTCAATGGCTGGCGCGCTGGCGGAGCG 347
Db 74 LeuGlnMetAlaTyrLysProValValValLeuAsnGlyLeuLysAlaValArgGluVal 93
QY 348 CTGGTGACCCACGCGGAGGACACCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407
Db 94 LeuValAsnCysGlyGluAspThrAlaAspArgProProValProIlePheAsnHisVal 113
QY 408 GGTGTTGGCGCGCTTCCAA----- 428
Db 114 GlyPheGlyHisAsnSerGlnGlyValAlaPheAlaArgTyrGlyProGlnTrpArgGlu 133
QY 428 ----- 428
Db 134 GlnArgArgPheCysValSerThrMetArgAspPheGlyValGlyLysSerLeuGlu 153
QY 429 -----GGAGCG 434
Db 154 GlnTrpValThrGluAlaGlyHisLeuCysAspAlaPheThrGlnGluAlaGlyHis 173
QY 435 CCCTTTCGCCCAACGCTCTTGTGACAAAGCGGTGAGCAACGCTGATGCTCCCTCCCTCACC 494
Db 174 ProPheAsnProThrThrLeuLeuAsnLysSerValCysAsnValIleSerSerLeuIle 193
QY 495 TCGCGGCGCGCTTCGAGTACGACGACCGCTCTCTCTCAGGCTCTGAGCCTAGCTCAG 554
Db 194 TyrAlaHisArgPheAspTyrGluAspProPheAsnSerLeuLeuLysMetLeuGln 213
QY 555 GAGGACTGAAGAGGAGTTCGGGCTTCTGCGCGAGGTGCTGAATGCTGTCTCCCTCCCTC 614
Db 214 GluSerPheGlyGluAspThrGlyPheIleAlaGluValLeuAsnAlaValProValLeu 233
QY 615 CTGCATATCCCGCGCTGGCTGGCAAGGTCTCTACGCTTCCAAAGCGCTTCTCTGACCCAG 674
Db 234 LeuArgIleProGlyLeuProGlyLysAlaPheProLysLeuThrAlaPheMetAspSer 253
QY 675 CTGGATGAGCTGAACCTGACGACAGGATGACCTGGGACCCAGCCAGCCCGCCCGAGAC 734
Db 254 LeuTyrLysMetLeuIleGluHisLysThrThrTrpAspProAlaGlnProProArgGly 273
QY 735 CTGACTGAGGCGCTTCTGCGCAGAGTGAAGGCAAGGGGAAACCTCTGAGAGCAGCTTC 794
Db 274 LeuThrAspAlaPheLeuAlaGluValGluLysAlaLysGlyArgProGluSerSerPhe 293
QY 795 AATGATGAGACCTGCGCATAGTGTGCTGCTGACCTGTCTCTGCGGAGATGTGACCCACC 854
Db 294 AsnAspGluAsnLeuArgMetValValAlaAspMetPheIleAlaGlyMetValThrThr 313


```
Db 223 IleProGluValLeuAuaThrPheProAlaLeuLeuAuaGileProGlyLeuAlaAspIys 242
QY 642 GTCTACGCTTCAAAAGGCTTCTGACGAGCTGATGCTCTAACTGAGCAGG 701
Db 243 ValPheGlnGlyGlnIysThrPheMetAlaLeuLeuAspLeuLeuAlaGluAsnArg 262
QY 702 ATGACCTGGGACCCAGCCAGCCCGGAGACCTGATGAGCGCTTCTCTGGCAGAGATG 761
Db 263 ThrThrTrpAspProAlaGlnProProArgAsnLeuThrAspAlaPheLeuAlaGluVal 282
QY 762 GAGAGGCCAAGGGAACCTGAGAGCAGCTTCATGATGAGAACCTGGCATAGTGTG 821
Db 283 GluIysAlaLysGlyAsnProGluSerPheAsnAspGluAsnLeuArgMetValVal 302
QY 822 GCTGACCTGTTCTGCGCGGATGTGACACCTGACGACGCTGCGCTGGCGCTCTCTG 881
Db 303 ValAspLeuPheThrAlaGlyMetValThrThrAlaThrThrLeuThrTrpAlaLeuLeu 322
QY 882 CTCATGATCTTACATCCGATGTGAGCGCGCTGTCCACAGGAGATCGACGCTGATA 941
Db 323 LeuMetIleLeuTyrrProAspValGlnArgValGlnGlnGluLeuAspGluValIle 342
QY 942 GGGAGGTGCGGCGACGACGAGATGGTGACACGCTCATGCTGCTGACACCTGCGCTG 1001
Db 343 GlyGlnValArgCysProGluMetThrAspGlnAlaHisMetProTyrrThrAsnAlaVal 362
QY 1002 ATTATGAGTGCAGCGCTTGGGAGACATCGTCCCGCTGGGTGTGACCATATGACATCC 1061
Db 363 IleHisGluValGlnArgPheGlyAspIleAlaProLeuAsnLeuProArgPheThrSer 382
QY 1062 COTGACATCAGATGACAGGCTTCCGCTATCCCTAAGGGAACGACACTCATCACCACTG 1121
Db 383 CysAspIleGluValGlnAspPheValIleProLysGlyThrLeuIleAsnLeu 402
QY 1122 TCATCGGTGCTGAGGATGAGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCGAGAAC 1181
Db 403 SerSerValLeuLysAspGluThrValTrpGluLysProHisArgPheHisProGluHis 422
QY 1182 TTCCTGGATGCCAGGCGCACCTTGTGAAGCGGAGGCTTCCGCTTCTTCTGACGAGC 1241
Db 423 PheLeuAspAlaGlnGlyAsnPheValLysHisGluAlaPheMetProPheSerAlaGly 442
QY 1242 CGCCGTGATGCTCGCGGAGCCCTGCGCGCATGAGGCTTCTTCTTCTTCTTCTTCTTCC 1301
Db 443 ArgArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrCys 462
QY 1302 CTGCTGAGCACTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1361
Db 463 LeuLeuGlnArgPheSerPheSerValProValGlnProArgProSerThrHisGly 482
QY 1362 GTCTTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1421
Db 483 PhePheAlaPheProValAlaProLeuProTyrrGlnLeuCysAlaValValArgGluGln 502
QY 1422 GGGGTA 1427
Db 503 GlyLeu 504
RESULT 14
ID CPD5 RAT STANDARD; PRT; 504 AA.
AC P12939;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D5 (EC 1.14.14.1) (CYP1D5) (P450-DB5) (P450-CWF1B)
DE (Debrisoquine 4-hydroxylase).
GN CYP2D5 OR CYP2D-5
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=90189185; PubMed=2107330;
RA Matsunaga E., Umeno M., Gonzalez F.J.;
RT "The rat P450 1D subfamily: complete sequences of four closely
RT linked genes and evidence that gene conversions maintained sequence
RT homogeneity at the heme-binding region of the cytochrome P450 active
RL J. Mol. Evol. 30:155-169(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90057430; PubMed=2819073;
RA Matsunaga T., Zanger U.M., Hardwick J.P., Gelboin H.V., Meyer U.A.,
RA Gonzalez F.J.;
RT "The CYP2D gene subfamily: analysis of the molecular basis of the
RT debrisoquine 4-hydroxylase deficiency in DA rats.";
RL Biochemistry 28:7349-7355(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89366685; PubMed=2771656;
RA Ishida N., Inuzuka C., Tawaragi Y., Sugita O., Nakazato H.,
RA Noguchi T., Sassa S., Kappas A.;
RT "Cytochrome P450CMF cDNA: nucleotide sequence of P450CMF1b.";
RL Nucleic Acids Res. 17:6407-6407(1989).
RN [4]
RP SEQUENCE OF 18-504 FROM N.A.
RX MEDLINE=89500091; PubMed=3190674;
RA Ishida N., Tawaragi Y., Inuzuka C., Sugita O., Kubota I.,
RA Nakazato H., Noguchi T., Sassa S.;
RT "Four species of cDNAs for cytochrome P450 isozymes immunorelated to
RT P450C-M/F encode for members of P450IID subfamily, increasing the
RL number of members within the subfamily.";
RL Biochem. Biophys. Res. Commun. 156:681-688(1988).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M25143; AAA41034.1; -
CC EMBL; X52030; CAA36272.1; -
CC EMBL; J02869; AAA41003.1; -
CC EMBL; M22329; AAA41045.1; -
CC PIR; S09611; O4RTD5.
CC HSSP; P00179; 1DT6.
CC InterPro; IPR001128; Cytochrome P450.
CC InterPro; IPR008069; EP450_CYP2D.
CC Pfam; PF00067; P450_1.
CC PRINTS; PR01686; EP450ICYP2D.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 446 446 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 504 AA; 57076 MW; F308EE06F2605DFE CRC64;
Alignment Scores: 9.25e-88 Length: 504
Pred. No.:
```

Score: 1616.50 Matches: 319
Percent Similarity: 72.12% Conservative: 38
Best Local Similarity: 64.44% Mismatches: 87
Query Match: 56.01% Indels: 51
DB: 1 Gaps: 1

US-09-820-788a-1 (1-1537) x CPD5_RAT (1-504)

QY 93 CTGGTCCCGCTGGCTGAGTGGCCATCTTCTGCTCTGCTGACCTGATGCAACCG 152
Db 9 LeuTriProMetAlaIlePheThrValIlePheIleLeuValAspLeuMetHisArg 28

QY 153 CGCCAAAGCTGGCTGACGCTACTCACCAGCGCCCTGCCACTGCGCGGCTGGGCAAC 212
Db 29 HisGlnArgTrpThrSerArgTrpProGlyProValProTrpProValLeuGlyAsn 48

QY 213 CTGCTGATGTGACTTCAGAACACACACATCTGCTTCGACAGTTGGCGCGCTTC 272
Db 49 LeuLeuGlnValAspProSerAsnMetProTyrSerMetTyrIleLeuGlnHisArgTyr 68

QY 273 GGGGACGTGTTACGCTGAGCTGGCTGGACCGCGTGTGTCTCAATGGCTGGCG 332
Db 69 GlyAspValPheSerLeuGlnMetGlyTyrLysProMetValIleValAsnArgLeuLys 88

QY 333 GCGTGGCGGCGCTGTGACCCACGGGAGGACACCGCGACCGCGCTGTGCGCC 392
Db 89 AlaValGlnGluValLeuValThrHisGlyGluAspThrAlaAspArgProValPro 108

QY 393 ATCACCAGATCTGCTGTTGGCGCGCTTCCTCAA----- 428
Db 109 IlePheLysCysLeuGlyValLysProArgSerGlnGlyValValPheAlaSerTyrGly 128

QY 428 ----- 428
Db 129 ProGluTrpArgGluGlnArgArgPheSerValSerThrLeuArgThrPheGlyMetGly 148

QY 428 ----- 428
Db 149 LysLysSerLeuGluGluTrpValThrLysGluAlaGlyHisLeuCysAspAlaPheThr 168

QY 429 -----GGAGCGCCCTTCGCCCCAACGGTCTCTGGACAAAGCCGTGAGCAACGG 479
Db 169 AlaGlnAsnGlyArgSerIleAsnProLysAlaMetLeuAsnLysAlaLeuCysAsnVal 188

QY 480 ATGCGCTCCCTCACCTGCGGCGCGCTTCGAGTACGACGACCTCGCTTCCTCAAGCTG 539
Db 189 IleAlaSerLeuIlePheAlaArgPheGluTyrGluAspProTyrLeuIleArgMet 208

QY 540 CTGGACCTAGCTCAGGAGGACTGAAGGAGGAGTGGGCTTCTGCGCGAGGTGTGAAT 599
Db 209 LeuThrLeuValGluGluSerLeuIleGluValSerGlyPheIleProGluValLeuAsn 228

QY 600 GCTGTCCCGCTCTCTGTCATATCCAGCGCTGGCTGGCAAGTCTCTACGCTTCCAAAG 659
Db 229 ThrPheProAlaLeuLeuArgIleProGlyLeuAlaAspLysValPheGlnGlyLys 248

QY 660 GCTTCTCCAGCAGCTGAGTGTCTAAGTACGACAGGACGCTGGACCCAGCC 719
Db 249 ThrPheMetAlaPheLeuAspAsnLeuLeuAlaGluAsnArgThrThrTrpAspProAla 268

QY 720 GAGCCCCCGGAGACTGAGCTGAGCGCTTCCTGGCAGAGATGAGAGCCAAAGGGGAC 779
Db 269 GlnProProArgAsnLeuThrAspAlaPheLeuAlaGluValGluLysAlaLysGlyAsn 288

QY 780 CCTGAGAGAGCTTCAATGATCAGACCTGCGCATAGTGTGGTGTCTCTCTGCTGCC 839
Db 289 ProGluSerSerPheAsnAspGluAsnLeuArgMetValValAspLeuPheThrAla 308

QY 840 GGGATGGTGCACCACTCGACCAAGCTGGCTGGCTGGCTGCTGCTCATGATCTTACATCG 899
Db 309 GlyMetValThrAlaThrThrLeuThrTrpAlaLeuLeuLeuMetIleLeuLysPro 328

QY 900 GATGTGACGCGCGTGTCCAACAGGAGATCGACGAGTGTGAGGCGAGGTGCGCGACCA 959

Db 329 AspValGlnArgArgValGlnGlnGluIleAspGluValIleGlyGlnValArgCysPro 348
QY 960 GAGATGGGTGACGAGCTCACATGCCCTACACACCTGCGTGTATTCATGAGGTGCGAGCG 1019
Db 349 GluMetThrAspGlnAlaHisMetProTyrThrAsnAlaValIleHisGluValGlnArg 368

QY 1020 TTTGGGACATCGTCCCTGCGGTGTGACCCATATGACATCCCGTGACATCGAAGTACAG 1079
Db 369 PheGlyAspIleAlaProLeuAsnLeuProArgIleThrSerCysAspIleGluValGln 388

QY 1080 GCTTCCGATCCCTAAGGAACGACACTCATCAACACTGTCTGCTGCTGCTGAAGGAT 1139
Db 389 AspPheValIleProLysGlyThrThrLeuIleIleAsnLeuSerSerValLeuLysAsp 408

QY 1140 GAGGCGCTCTGGGAGAGCCCTTCGCTTCCACCCCGAACACTCTCTGATGATGCCAGGGC 1199
Db 409 GluThrValTrpGluLysProLeuArgPheHisProGluHisPheLeuAspAlaGlnGly 428

QY 1200 CACTTTGTGAAGCCGAGGCTTCCTGCTTCTCAGCAGGCGCGCTGCTGCTGCTCGGG 1259
Db 429 AsnPheValLysHisGluAlaPheMetProPheSerAlaGlyArgAlaCysLeuGly 448

QY 1260 GAGCCCTGCGCGGATGAGCTTCTCTTCTTCTCCTCCTGCTGCTGCTGCTGCTGCTGCT 1319
Db 449 GluProLeuAlaArgMetGluLeuPheLeuPheThrCysLeuLeuGlnHisPheSer 468

QY 1320 TTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1379
Db 469 PheSerValProAlaGlyGlnProArgProSerThrLeuGlyAsnPheAlaIleSerVal 488

QY 1380 ACCCATCCCTATCAGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1424
Db 489 AlaProLeuProTyrGlnLeuCysAlaAlaValArgGluGlnGly 503

RESULT 15
CPDG_CAVPO STANDARD; PRT; 500 AA.

ID CPDG_CAVPO STANDARD; PRT; 500 AA.
AC Q64403; O54866;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D16 (EC 1.14.14.1) (CYP2D16).
GN CYP2D16.

OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-38.
RC STRAIN=13; TISSUE=Adrenal cortex;
RX MEDLINE=95251703; PubMed=773969;
RA Jiang Q., Voigt J.M., Colby H.D.;
RT "Molecular cloning and sequencing of a guinea pig cytochrome P4502D
(CYP2D16): high level expression in adrenal microsomes.";
RL Biochem. Biophys. Res. Commun. 209:1149-1156(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley white; TISSUE=adrenal gland;
RA Sun Y., Voigt J.M., Pierce J.C., Colby H.D.;
RT "The gene sequence of a xenobiotic metabolism-related cytochrome P450
isozyme (CYP2D16) in guinea pig adrenal gland.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
monooxygenases. In liver microsomes, this enzyme is involved in an
NADPH-dependent electron transport pathway. It oxidizes a variety
of structurally unrelated compounds, including steroids, fatty
acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE INNER ZONE OF
THE ADRENAL CORTEX.

CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U21486; AA68479.1; -.
 DR EMBL; AF020345; AAB94568.1; -.
 DR PIR; JC4153; JC4153.
 DR HSP; P00179; 1D76.
 DR InterPro; IPR001128; Cytochrome P450.
 DR InterPro; IPR008069; EP450_CYP2D.
 DR Pfam; PF00067; p450; 1.
 DR PRINTS; PR01686; EP450ICYP2D.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT DOMAIN 81 84 POLY-VL.
 FT METAL 446 446 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT CONFLICT 123 123 I -> V (IN REF. 2).
 FT CONFLICT 127 127 Y -> N (IN REF. 2).
 FT CONFLICT 148 148 G -> R (IN REF. 2).
 SQ SEQUENCE 500 AA; 55800 MW; 2429247E49BF6B24 CRC64;

Alignment Scores:

Pred. No.: 1,94e-87 Length: 500
 Score: 1611.00 Matches: 317
 Percent Similarity: 72.20% Conservative: 44
 Best Local Similarity: 63.40% Mismatches: 85
 Query Match: 55.82% Indels: 54
 DB: 1 Gaps: 2

US-09-820-788a-1 (1-1537) x CPDQ_CAVPO (1-500)

QY 78 ATGGGGCTA-----GAGCAGCTGGTGGCCCTGGCGGTGATGATGGCCATCTTCCTG 128
 Db 1 MetGlyLeuLeuThrGlyAspAlaLeuPheSerValAlaValAlaValAlaLeuPheLeu 20
 QY 129 CTCCTGGTGGACCTGATCAGCCGGCCAAAGCTGGGTGACGCTACTCACCAGGCCCC 188
 Db 21 LeuLeuValAspLeuMetHisArgArgGlnArgTrpAlaAlaArgTrpProGlyPro 40
 QY 189 CTGCACCTGCCGGCTGGCAACCTGCTGTGATGTGACCTTCCAGAACACACCATACTGC 248
 Db 41 ValProValProGlyLeuGlyAsnLeuLeuGlnValAspPheGluAsnMetAlaTrpSer 60
 QY 249 TTCGACAGTGGCGCGCGCTTGGGAGCAGTGTTCAGCTGCAGCTGGCGCTGGAGCGCG 308
 Db 61 CysAspLysLeuArgHisGlnPheGlyAspValPheSerLeuGlnPheValTrpThrPro 80
 QY 309 GTGGTCGTCTCAATGGGTGGCGCTGGCGAGGCGCTGGTGCACCCAGCGGAGGAC 368
 Db 81 ValValValValAsnGlyLeuLeuAlaValArgGluAlaLeuValAsnAsnSerThrAsp 100
 QY 369 ACCGCGGACCGCGCCCTGTGCCATCACCAGATCCTGGGTGGTGGCGCGCTTCCCAA 428
 Db 101 ThrSerAspArgProThrLeuProThrAsnAlaLeuLeuGlyPheGlyProLysAlaGln 120
 QY 428 ----- 428
 Db 121 GlyValIleGlyAlaTrpTrpGlyProAlaTrpArgGluGlnArgArgPheSerValSer 140
 QY 428 ----- 428
 Db 141 SerLeuArgAsnPheGlyLeuGlyLysLysSerLeuGluGlnTrpValThrGluAla 160
 QY 429 -----GGACGCCCTTTCGCCCAACGGTCTC 455
 |||:|||||

Db 161 AlaCysLeuCysAlaAlaPheThrAsnHisAlaGlyGlnProPheCysProLysAlaLeu 180
 QY 456 TTGGACAAGCCGGAGCAAGTGTGATCGCTCCCTCCCTACCTGCGGGCGCGCTTCGAGTAC 515
 Db 181 LeuAsnLysAlaValCysAsnValIleSerSerLeuLeuTyrAlaAlaArgPheAspTrp 200
 QY 516 GACCAACCTCGCTTCCTCAGGCTCTGAGACCTAGCTCAGGAGGAGCTGAAGAGGAGTCG 575
 Db 201 AspAspProMetValLeuArgLeuLeuGluPheLeuGluGluThrLeuArgGluAsnSer 220
 QY 576 GGCTTTCTGGCGAGGTGCTGAATGTGTCCTCCCTCTCTGTCATATCCCAAGCGCTGCT 635
 Db 221 SerLeuLysIleGlnValLeuAsnSerIleProLeuLeuLeuArgIleProCysValAla 240
 QY 636 GGCAGGTCTCTACGCTTCCAAAAGGCTTCTGACCCAGCTGGATGAGCTGCTAACTGAG 695
 Db 241 AlalysValLeuSerAlaGlnArgSerPheIleAlaLeuAsnAspLysLeuLeuAlaGlu 260
 QY 696 CACAGGATGACCTGGGACCCAGCCCGAGCCCTGAGCTGACTGAGGCTTCTCTGGCA 755
 Db 261 HisAsnThrGlyTrpAlaProAspGlnProProArgAspLeuThrAspAlaPheLeuThr 280
 QY 756 GAGATGAGAGGCAAGGGGAACCTGAGAGAGCTTCAATGATGAGAACTGCGCAT 815
 Db 281 GluMetHisLysAlaGlnGlyAsnSerGluSerSerPheAsnAspGluAsnLeuArgLeu 300
 QY 816 GTGGTGGCTGACCTGTTCTCTGCGGGATGTGTGACCACTGACCACTGCGCTGGCGGC 875
 Db 301 LeuValSerAspLeuPheGlyAlaGlyMetValThrSerValThrLeuSerTrpAla 320
 QY 876 CTCCTGCTCATGATCTACATCCGATGTCGAGCGCGCTGTCACACAGGAGATCGACGAC 935
 Db 321 LeuLeuLeuMetIleLeuHisProAspValGlnArgHisValGlnGluLeuAspGlu 340
 QY 936 GTGATAGGGAGGTGCGCGACACAGAGATGGGTGACAGGCTCATGCTCCACTACCACT 995
 Db 341 ValIleGlyGlnValArgCysProGluMetAlaAspGlnAlaHisMetProPheThrAsn 360
 QY 996 GCGTGTATCATGAGGTGCAGCGCTTTGGGACATCGTCCCTGGGTGTGACCCATATG 1055
 Db 361 AlaValIleHisGluValGlnArgPheAlaAspIleValProMetGlyValProHisMet 380
 QY 1056 ACATCCCGTGACATCGAAGTACAGGGCTTCGCGCATCCCTAAGGGAACGACACTCATACC 1115
 Db 381 ThrSerArgAspThrGluValGlnGlyPheLeuLeuProLysGlyThrMetLeuPheThr 400
 QY 1116 AACCTGTCATCGTGTGAGAGATGAGCGCTGTGGGAGAGAGCCCTTCGCTTCCACCCC 1175
 Db 401 AsnLeuSerSerValLeuLysAspGluThrValTrpGluLysProLeuHisPheHisPro 420
 QY 1176 GAACACTTCTGGATGCCAGGGCCACTTGTGTGAGCGGAGGCTTCTGCTTCTTCTCA 1235
 Db 421 GlyHisPheLeuAspAlaGluGlyArgPheValLysArgGluAlaPheMetProPheSer 440
 QY 1236 GCAGCGCGCTGATGCTCGGGAGCCCTGCGCGAGCCCTGCGCGCATGAGGCTTCTTCTTTC 1295
 Db 441 AlaGlyProArgIleCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePhe 460
 QY 1296 ACCTCCCTGTGACGACTTCAGCTTCTCGGTGGCCCACTGGACAGCCCGCCGCGCCAC 1355
 Db 461 ThrSerLeuLeuGlnArgPheSerPheSerValProGluGlyGlnProArgProSerAsp 480
 QY 1356 CATGGTCTTTGCTTCTTCCTGTCACCCCATCCCTATGAGCTTGTGCTGCTGCCCGC 1415
 Db 481 ArgGlyAlaProTyrLeuValValLeuProSerProTyrGlnLeuCysAlaValLeuArg 500

Search completed: February 25, 2004, 02:37:39
 Job time : 34.7487 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 03:04:11 ; Search time 59 Seconds
(without alignments)
2135.868 Million cell updates/sec

Title: US-09-820-788A-2

Perfect score: 2330

Sequence: 1 MGLEALVPLAIVAIALLLV.....HGVFAFLVTPSPFELCAVPR 446

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2330	100.0	446	AAO26404	Human drp
2	2290.5	98.3	497	ADb25834	Human CYP
3	2289.5	98.3	497	Aar72376	Human aux
4	2289.5	98.3	497	Aar93183	Human Cyt
5	2288.5	98.1	497	ADb25832	Human CYP
6	2286.5	98.1	497	ABU09598	Human CYP
7	2286.5	98.1	497	ADb25831	Human CYP
8	2285.5	98.1	497	Aar72377	Human aux
9	2285.5	98.1	497	Aar93184	Human Cyt
10	2285.5	98.1	497	Aar81462	Human CYP
11	2281.5	97.9	497	ABU09563	Human CYP
12	2285.5	98.1	497	AAO26405	Human CYP
13	2285.5	98.1	497	ADb25833	Human CYP
14	2282.5	98.0	497	ABU09593	Human CYP
15	2282.5	98.0	497	ABR82026	Human CYP
16	2282.5	98.0	497	ADb60558	Human CYP
17	2281.5	97.9	497	Aar72375	Human aux
18	2281.5	97.9	497	Aar93182	Human CYP
19	2279.5	97.8	497	AAW44869	Cytochrom
20	2277.5	97.7	497	Aar72378	Human aux
21	2277.5	97.7	497	Aar93185	Human CYP
22	2263	97.1	502	ABU09595	Human CYP
23	2259	97.0	502	ABU09594	Human CYP
24	2093.5	89.8	497	AAE05171	Human drp
25	1708.5	73.3	500	ADb60556	Rat Prote

26	1614.5	69.3	504	6	AAO22644	Protein o
27	1011	43.4	195	2	AAW44870	Cytochrom
28	1008	43.3	436	4	AAU68579	Human nov
29	890	38.2	184	4	AAQ74361	Human col
30	857.5	36.8	497	5	AAU91321	Killifish
31	833.5	35.8	502	7	ADD46231	Human Pro
32	828.5	35.6	502	7	ADD46229	Rat Prote
33	821.5	35.3	497	5	AAU91322	Killifish
34	801	34.4	494	7	ADb63662	Rat Prote
35	797.5	34.2	490	2	AAR72370	Human aux
36	797.5	34.2	490	2	AAR89861	Cytochrom
37	797.5	34.2	490	2	AAR93177	Human CYP
38	797.5	34.2	490	2	AAR81463	Human CYP
39	797.5	34.2	490	2	AAW64072	Human CYP
40	797.5	34.2	490	6	ABU96553	Human CYP
41	796.5	34.2	490	6	ABU09971	Human CYP
42	795.5	34.1	490	2	AAR72372	Human aux
43	795.5	34.1	490	2	AAR93179	Human CYP
44	794.5	34.1	490	2	AAR72371	Human aux
45	794.5	34.1	490	2	AAR93178	Human CYP

ALIGNMENTS

RESULT 1
AAO26404
ID AAO26404 standard; protein; 446 AA.
XX AC AAO26404;
XX AC
XX 30-JAN-2003 (first entry)
XX DT
XX DE Human drug-metabolising enzyme protein.
XX KW Cytostatic; antiparkinsonian; gene therapy; cancer; Parkinson's disease;
XX KW human drug-metabolising protein; enzyme.
XX OS Homo sapiens.
XX PN WO200279233-A1.
XX PD 10-OCT-2002.
XX PF 01-APR-2002; 2002WO-US009738.
XX PR 30-MAR-2001; 2001US-00820788.
XX PA (PEKE) PE CORP NY.
XX PA (DFRA/) DI FRANCESCO V.
XX PA (BEAS/) BEASLEY E M.
XX PI Shao W, Yan C;
XX DR WPI; 2003-040649/03.
XX DR N-PSDB; AAL53565, AAL53566.
XX PT New human drug-metabolizing proteins and nucleic acids related to the
XX PT Cytochrome P450 IID drug-metabolizing enzyme subfamily, useful for
XX PT treating a condition mediated by a human enzyme protein e.g., cancer.
XX PS Claim 1; Fig 2A; 72pp; English.
XX CC The invention relates to a novel isolated polypeptide comprising a 446-
XX CC amino acid sequence or its allelic variant, orthologue or fragment. The
XX CC allelic variant or orthologue is encoded by a nucleic acid that
XX CC hybridises under stringent conditions to the opposite strand of the
XX CC nucleic acid comprising a sequence having 1537 or 10278 bp. The fragment
XX CC comprises at least 10 contiguous amino acids of the 446-amino acid
XX CC sequence. The polypeptide is useful for preparing a pharmaceutical
XX CC composition for treating a disease or condition mediated by a human
XX CC enzyme protein, e.g. cancer or Parkinson's disease. This sequence
XX CC represents the human drug-metabolising protein of the invention

```

XX SQ Sequence 446 AA;
Query Match 100.0%; Score 2330; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.3e-223;
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLEALVPLAVIAI FLLLLVDMHRRORWAARYSPGLPLGLGNLLHVDQNTPYCFDQ 60
DB 1 MGLEALVPLAVIAI FLLLLVDMHRRORWAARYSPGLPLGLGNLLHVDQNTPYCFDQ 60
QY 61 LRRRFCDVFSQLAWTPVVLNGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQGRP 120
DB 61 LRRRFCDVFSQLAWTPVVLNGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQGRP 120
QY 121 FRPGLLDKAVSNVIASLTGRRPFYDDPRFLRLDLAQEGIKESGFLREVNLAVPVLL 180
DB 121 FRPGLLDKAVSNVIASLTGRRPFYDDPRFLRLDLAQEGIKESGFLREVNLAVPVLL 180
QY 181 HIPALAGKVLRFQKAFITQDDELLTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPSSFN 240
DB 181 HIPALAGKVLRFQKAFITQDDELLTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPSSFN 240
QY 241 DENLRIVVADLFSAGWTTTSTTLAWGLLMLILHPDVQRRVQOEIDDVICQVRRPEMGDOA 300
DB 241 DENLRIVVADLFSAGWTTTSTTLAWGLLMLILHPDVQRRVQOEIDDVICQVRRPEMGDOA 300
QY 301 HMPYTTAVIHEVQRFQDIPVGLVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEK 360
DB 301 HMPYTTAVIHEVQRFQDIPVGLVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEK 360
QY 361 PRRFPEHFLDAQGHFVPEAFPLPSAGRRACLGEPLARMELFLFTSLLOHFSFVPTG 420
DB 361 PRRFPEHFLDAQGHFVPEAFPLPSAGRRACLGEPLARMELFLFTSLLOHFSFVPTG 420
QY 421 QPRSHHGVAFLVTPSPVELCAVPR 446
DB 421 QPRSHHGVAFLVTPSPVELCAVPR 446

RESULT 2
ADB25834
ID ADB25834 standard; protein; 497 AA.
XX AC ADB25834;
XX DT 20-NOV-2003 (first entry)
XX DE Human CYP2D6-related protein #4.
XX KW human; mutant CYP2D6 gene; drug analysis; drug testing.
XX OS Homo sapiens.
XX PN WO2003050282-A1.
XX PD 19-JUN-2003.
XX PF 05-DEC-2002; 2002WO-JP012748.
XX PR 06-DEC-2001; 2001JP-00372548.
XX PA (TSUR ) TSUMURA & CO.
XX PI Taniyama M, Ogawa K, Tsuchiya N, Hibino T;
XX DR WPI: 2003-505401/47.
XX DR N-PSDB; ADB25779.
XX PT Genetic polymorphisms of CYP2D6 gene in human population for analysis of
XX PT drug effect on individual patients and testing of new drugs.
XX PS Claim 8; Page 50-54; 75pp; Japanese.

XX The invention comprises mutant forms of the human CYP2D6 gene, containing
CC one or more of the following mutations G125A, C1858T, T2874C and C2875T.
CC The mutant human CYP2D6 genes of the invention are useful for analysing
CC the effect of drugs on individual patients and testing of new drugs. The
CC present amino acid sequence represents a human protein of the invention.
XX SQ Sequence 497 AA;
Query Match 98.3%; Score 2290.5; DB 6; Length 497;
Best Local Similarity 89.5%; Pred. No. 1.3e-219;
Matches 445; Conservative 1; Mismatches 0; Indels 51; Gaps 1;

QY 1 MGLEALVPLAVIAI FLLLLVDMHRRORWAARYSPGLPLGLGNLLHVDQNTPYCFDQ 60
DB 1 MGLEALVPLAVIAI FLLLLVDMHRRORWAARYSPGLPLGLGNLLHVDQNTPYCFDQ 60
QY 61 LRRRFCDVFSQLAWTPVVLNGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQ--- 117
DB 61 LRRRFCDVFSQLAWTPVVLNGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQGVF 120
QY 118 -----GRFRPNGLLIDK 129
DB 121 LARYCPAWREQRFSVSTLRNLGLGKKSLEQWVTEEAACLAAPANHSRGRFRPNGLLIDK 180
QY 130 AVSNVIASLTGRRPFYDDPRFLRLDLAQEGIKESGFLREVNLAVPVLLHIIPALAGKV 189
DB 181 AVSNVIASLTGRRPFYDDPRFLRLDLAQEGIKESGFLREVNLAVPVLLHIIPALAGKV 240
QY 190 LRFQKAFITQDDELLTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPSSFNENLRIVVA 249
DB 241 LRFQKAFITQDDELLTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPSSFNENLRIVVA 300
QY 250 DLFSAGWTTSTTLAWGLLMLILHPDVQRRVQOEIDDVIGOVRRPEMGDOAHMPTTAVI 309
DB 301 DLFSAGWTTSTTLAWGLLMLILHPDVQRRVQOEIDDVIGOVRRPEMGDOAHMPTTAVI 360
QY 310 HEVQRFQDIPVGLVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEKPFHFPHF 369
DB 361 HEVQRFQDIPVGLVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEKPFHFPHF 420
QY 370 LDAQGHFVPEAFPLPSAGRRACLGEPLARMELFLFTSLLOHFSFVPTGQPRPSHHGV 429
DB 421 LDAQGHFVPEAFPLPSAGRRACLGEPLARMELFLFTSLLOHFSFVPTGQPRPSHHGV 480
QY 430 FAFLVTPSPVELCAVPR 446
DB 481 FAFLVTPSPVELCAVPR 497

RESULT 3
AAR72376
ID AAR72376 standard; protein; 497 AA.
XX AC AAR72376;
XX DT 25-MAR-2003 (revised)
XX DT 15-NOV-1995 (first entry)
XX DE Human auxillary cytochrome P450 species 2D6 variant 1 protein.
XX KW Human cytochrome P450; amplification; PCR; primer; expression vector;
XX KW yeast NADPH-P450 reductase; safety; fusion protein; metabolite;
XX KW carcinogen; mutagen; liver metabolism.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 296
XX FT /note= "Cys to Arg variation"
XX PN EP644267-A2.

```

PD 22-MAR-1995.
XX
PF 20-JUL-1994; 94EP-00111298.
XX
XX 20-JUL-1993; 93JP-00201120.
PR 21-JUL-1993; 93JP-00180246.
PR 30-JUL-1993; 93JP-00208279.
XX
XX (SUMO) SUMITOMO CHEM CO LTD.
XX
XX Hayaashi K, Sakaki T, Yabusaki Y, Komai K, Kaneko H, Nakatsuka I;
XX WPI; 1995-116991/16.
XX N-PSDB; AAQ87730.
XX
XX Evaluation of safety of a chemical cpd. - using recombinant yeast
PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase.
XX
XX Example; Page 87-89; 124pp; English.
XX
XX The amino acid sequence of the human auxillary cytochrome P450 species
CC 2D6 variant 1. This variant contains a variation at residue 296: Cys to
CC Arg, caused by a variation at base 886: T to C in the DNA sequence. The
CC CDNA was amplified by PCR using the primers AAQ87763-6. The product was
CC cloned into the yeast expression vectors pAAH5N or pAHR to produce the
CC vectors p2D6 variant 1 for the expression of the cytochrome P450 alone or
CC p2D6R variant 1 for co-expression with the yeast NADPH-P450 reductase.
CC The vectors are used in a method for evaluating the safety of a chemical
CC compound by reacting the chemical compound with recombinantly produced
CC human cytochrome P450 molecular species 1A2 (AAQ87714), 2C9 (AAQ87715),
CC 2E1 (AAQ87716), or 3A4 (AAQ87718), or their auxillary species and
CC variants (AAQ87718-32), and yeast NADPH-P450 reductase, either as a fused
CC protein or in cell extracts, and analysing the resulting metabolite to
CC assess the safety of the chemical compound. The method is useful for
CC determining whether the chemical compound, or its metabolite, will be
CC converted into a carcinogenic or mutagenic form through metabolism in the
CC liver. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 497 AA;
SQ

Query Match 98.3%; Score 2289.5; DB 2; Length 497;
Best Local Similarity 89.5%; Pred. No. 1.7e-219;
Matches 445; Conservative 0; Mismatches 1; Indels 51; Gaps 1;
QY 1 MGLEALVPLAVIVAIFFLLVLMHRRQWAAARYSGPLPLGGLNLLHVDFTQNTYCFDQ 60
DB 1 MGLEALVPLAVIVAIFFLLVLMHRRQWAAARYSGPLPLGGLNLLHVDFTQNTYCFDQ 60
QY 61 LRRRFGDVFSLQAWTPVVVNLGLAAVREALVTHGEDTADRPVPIITQLGFGPRSQ--- 117
DB 61 LRRRFGDVFSLQAWTPVVVNLGLAAVREALVTHGEDTADRPVPIITQLGFGPRSQGVF 120
QY 118 -----GRPRPNGLLDK 129
DB 121 LARYGPAMREQRFSVSTLRNLGLKKSLEQWVTEAACLCFAAFNHSGRPPRPNGLLDK 180
QY 130 AVSNVIASITCGRRREYDDPRFLRLDLDAQGLKEESGFLREVNAPVLLHHPALAGKV 189
DB 181 AVSNVIASITCGRRREYDDPRFLRLDLDAQGLKEESGFLREVNAPVLLHHPALAGKV 240
QY 190 LRFQKAFITQLDELTEHMTWDPAPPRDLTEAFLEAMEKAKGNPSSFNENLRIVVA 249
DB 241 LRFQKAFITQLDELTEHMTWDPAPPRDLTEAFLEAMEKAKGNPSSFNENLRIVVA 300
QY 250 DLFSAGMVTSTTLAWGLLMLLHPDVQRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 309
DB 301 DLFSAGMVTSTTLAWGLLMLLHPDVQRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 360
QY 310 HEVQRFGDVLPGVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEKPFRRPHFPH 369
DB 361 HEVQRFGDVLPGVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEKPFRRPHFPH 420
QY 370 LDAQGHFVKPEAFLPFSAGRRACLGELPLARMELFLFTSLLOHFSFVPTGQPRPSHHGV 429

DB 421 LDAQGHFVKPEAFLPFSAGRRACLGELPLARMELFLFTSLLOHFSFVPTGQPRPSHHGV 480
QY 430 FAFLVTPSPYELCAVPR 446
DB 481 FAFLVTPSPYELCAVPR 497
RESULT 4
AAR93183
ID AAR93183 standard; protein; 497 AA.
XX AC AAR93183;
XX 11-OCT-1996 (first entry)
XX Human cytochrome P450 molecular species 2D6 variant #2 protein.
XX Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;
KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;
KW evaluation; safety; fusion protein; metabolite; detoxification;
XX carcinogenic.
XX Homo sapiens.
XX JP08056695-A.
XX 05-MAR-1996.
XX 15-JUL-1994; 94JP-00164184.
XX 20-JUL-1993; 93JP-00201120.
XX 30-JUL-1993; 93JP-00208279.
XX 17-JUN-1994; 94JP-00136053.
XX (SUMO) SUMITOMO CHEM CO LTD.
XX WPI; 1996-182311/19.
XX N-PSDB; AAT28396.
XX
XX Novel method for the evaluation of the safety of a cpd. - using a human
PT cytochrome P450 and yeast NADPH reductase to determine whether the
PT analyte cpd. is detoxified or metabolised to a carcinogen.
XX
XX Example 1; Page 51-53; 74pp; Japanese.
XX This is the amino acid sequence of the human cytochrome P450 molecular
CC species 2D6 variant #2 protein. The corresp. gene was amplified from a
CC human liver derived cDNA library as 2 fragments of 0.4 and 0.9 kb using
CC primers AAT26953-6. The prod. was cloned into the yeast expression vector
CC pAAH5N to generate plasmid p2D6 for prodn. of the cytochrome only or into
CC the vector pAHR to generate the plasmid p2D6R for co-prodn. with the
CC yeast NADPH-P450 reductase. The sequence is placed under control of the
CC yeast ADH gene promoter and terminator. The vectors are used in a method
CC for evaluating the safety of a cpd. by reacting the test cpd. with
CC recombinantly produced human cytochrome P450 mol. species 1A2 (AAT28380),
CC 2C9 (AAT28381), 2E1 (AAT28382), 3A4 (AAT28383) or their variants
CC (AAT28384-98) together with yeast NADPH-P450 reductase (either as a fused
CC protein or as a cell extract) and analysing the resultant metabolite. The
CC cpd. is considered "safe" if it is detoxified or not rendered
CC carcinogenic or "unsafe" if it is not detoxified or is metabolised to a
CC carcinogenic cpd
XX Sequence 497 AA;
SQ

Query Match 98.3%; Score 2289.5; DB 2; Length 497;
Best Local Similarity 89.5%; Pred. No. 1.7e-219;
Matches 445; Conservative 0; Mismatches 1; Indels 51; Gaps 1;
QY 1 MGLEALVPLAVIVAIFFLLVLMHRRQWAAARYSGPLPLGGLNLLHVDFTQNTYCFDQ 60
DB 1 MGLEALVPLAVIVAIFFLLVLMHRRQWAAARYSGPLPLGGLNLLHVDFTQNTYCFDQ 60

Qy	61	LRRRFGDVFSIQLAWTPVVVWNLGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQV	117
Db	61	LRRRFGDVFSIQLAWTPVVVWNLGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQGVF	120
Qy	118	-----GRPRRPNGLLDK	129
Db	121	LARYCPAWREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAPAFHNSGRPRPNGLLDK	180
Qy	130	AVSNVIASLTCRRRFEYDDPRFLRLDLLDAQEGLKEESGFLREVLNAVPPVLLHI PALAGKV	189
Db	181	AVSNVIASLTCRRRFEYDDPRFLRLDLLAQEGLKEESGFLREVLNAVPPVLLHI PALAGKV	240
Qy	190	LRFOKAFULTQDELLETHRMTWDPQAPPRDLTEAFLAEMEKAKGNPESFNENDENLRIVVA	249
Db	241	LRFOKAFULTQDELLETHRMTWDPQAPPRDLTEAFLAEMEKAKGNPESFNENDENLRIVVA	300
Qy	250	DLFSAGMVTTSITTLAWGLLLMLHPDVQRRVQOEIDDDVIGQVRRPEMGDQAHMPTTVAVI	309
Db	301	DLFSAGMVTTSITTLAWGLLLMLHPDVQRRVQOEIDDDVIGQVRRPEMGDQAHMPTTVAVI	360
Qy	310	HEVQRFGBIVPLGVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEKDFRPHPEHF	369
Db	361	HEVQRFGBIVPLGVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEKDFRPHPEHF	420
Qy	370	LDAQGHFVKPFAFLPFSAGRACCLGEPLARMELFUFTSLLOHFSFSVPTGQPRSSHGV	429
Db	421	LDAQGHFVKPFAFLPFSAGRACCLGEPLARMELFUFTSLLOHFSFSVPTGQPRSSHGV	480
Qy	430	FAFLVTPSPPYELCAVPR	446
Db	481	FAFLVTPSPPYELCAVPR	497

RESULT 5	
ADB25832	
ID	ADB25832 standard; protein; 497 AA.
XX	
AC	ADB25832;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Human CYP2D6-related protein #2.
XX	
KW	human; mutant CYP2D6 gene; drug analysis; drug testing.
XX	
OS	Homo sapiens.
XX	
OS	
XX	
PN	WO2003050282-A1.
XX	
PD	19-JUN-2003.
XX	
XX	
PF	05-DEC-2002; 2002WO-JP012748.
XX	
XX	
PR	06-DEC-2001; 2001JP-00372548.
XX	
PA	(TSUR) TSUMURA & CO.
XX	
PI	Taniyama M, Ogawa K, Tsuchiya N, Hibino T;
XX	
DR	WPI; 2003-505401/47.
DR	N-ESDB; ADB25777.
XX	
PT	Genetic polymorphisms of CYP2D6 gene in human population for analysis of
PT	drug effect on individual patients and testing of new drugs.
XX	
PS	Claim 8; Page 43-46; 75pp; Japanese.
XX	
CC	The invention comprises mutant forms of the human CYP2D6 gene, containing
CC	one or more of the following mutations G125A, C1858T, T2874C and C3875T.
CC	The mutant human CYP2D6 genes of the invention are useful for analysing
CC	the effect of drugs on individual patients and testing of new drugs. The
CC	present amino acid sequence represents a human protein of the invention.
XX	

XX Milos PM, Webb SM;
 XX WPI; 2003-373769/36.
 DR N-PSDB; ACA61331.
 XX
 XX New cytochrome P450 2D6 gene variants and polypeptides, useful for
 PT determining if a subject has or is at risk of developing a drug
 PT sensitivity condition or disorder that is associated with an aberrant
 PT CYP2D6 activity.
 XX
 XX Claim 4; Fig 7; 88pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising a cytochrome
 CC P450 2D6 gene variant, e.g. G5799C or C5816AT (referring to the genomic
 CC sequence or the same variant nucleotide in the corresponding cDNA
 CC sequences). Also included are probes, primers (allele specific
 CC oligonucleotides) and arrays used to detect and/or amplify the CYP2D6
 CC gene polymorphic regions, the variant polypeptides, antibodies which are
 CC capable of distinguishing between the variant and wild-type polypeptides,
 CC determining whether a subject has a genetic deficiency for metabolising a
 CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
 CC determining whether an individual is susceptible to being a poor
 CC metaboliser of drugs. The DNA probe is useful for hybridising to a
 CC variant form of the CYP2D6 gene. The primer is useful for amplifying the
 CC C5816TA allelic variant. The allele specific nucleotide is useful for the
 CC detection of the C5816TA allelic variant. The methods are useful for
 CC determining whether a subject has a genetic deficiency for metabolising a
 CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and
 CC determining if an individual is susceptible to being a poor metaboliser
 CC of drugs. The nucleic acids are useful as probes or primers for
 CC determining whether a subject has a genetic deficiency for metabolising
 CC drugs that are substrates of P450 CYP2D6. The methods are useful for
 CC determining if a subject has or is at risk of developing a drug
 CC sensitivity condition or disorder that is associated with an aberrant
 CC CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
 CC aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
 CC appropriate drugs or determining the course of treatment to administer to
 CC a subject to treat cardiovascular or psychiatric disorders, or for
 CC treating a subject with a drug sensitivity or disorder associated with a
 CC specific allelic variant of a polymorphic region of the CYP2D6 gene. The
 CC antibodies are useful for monitoring CYP2D6 protein levels in an
 CC individual for determining whether a subject has a disease or conditions
 CC associated with an aberrant CYP2D6 protein level. The gene is located on
 CC human chromosome 22. The present sequence is the G5799C variant CYP2D6
 CC protein
 XX
 SQ Sequence 497 AA;

Query Match 98.1%; Score 2286.5; DB 6; Length 497;
 Best Local Similarity 89.3%; Pred. No. 3.3e-219;
 Matches 444; Conservative 1; Mismatches 1; Indels 51; Gaps 1;

QY 1 MGLEALVPLAVIVAIFLLVLDLMHRRQWRWAARYSGPLPLGGLNLLHVDFTQNTYCFDQ 60
 DB 1 MGLEALVPLAVIVAIFLLVLDLMHRRQWRWAARYSGPLPLGGLNLLHVDFTQNTYCFDQ 60
 QY 61 LRRFGDVFSLOLAWTPVWVNLGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQ--- 117
 DB 61 LRRFGDVFSLOLAWTPVWVNLGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQGVF 120
 QY 118 -----GRPPRPNGLLDDK 129
 DB 121 LARYGPAWRQRRFVSITLRNLGLGKSLQWVTEAAACLAFAHSGRPPRPNGLLDDK 180
 QY 130 AVSNVIASITCGRRFYDDPRFLRLDLAQEGKKESGFLREVNAVVPVLLHIPALAGKV 189
 DB 181 AVSNVIASITCGRRFYDDPRFLRLDLAQEGKKESGFLREVNAVVPVLLHIPALAGKV 240
 QY 190 LRFQKAFITQLODELLTEHMTWDPAPQPPDLTEAFIAEMEKAKGNPSSFNDENIRIYVA 249
 DB 241 LRFQKAFITQLODELLTEHMTWDPAPQPPDLTEAFIAEMEKAKGNPSSFNDENIRIYVA 300

QY 250 DLFSAGMVTSTTTLAWGLLLMLILHDPVQRRVQOEIDDDVIGQVRRPEMGDQAHPYTTAVI 309
 DB 301 DLFSAGMVTSTTTLAWGLLLMLILHDPVQRRVQOEIDDDVIGQVRRPEMGDQAHPYTTAVI 360
 QY 310 HEVQRFSGDIVPLGVTHMTSRDIEVQGFRI PKGTTTLITNLSSVLKDEAVWEKPRFHPHF 369
 DB 361 HEVQRFSGDIVPLGVTHMTSRDIEVQGFRI PKGTTTLITNLSSVLKDEAVWEKPRFHPHF 420
 QY 370 LDAQGHFVKPEAFIPFSSAGRRACLGELPLARMELFLFTSLQHFQSFVPTGQPRSHGV 429
 DB 421 LDAQGHFVKPEAFIPFSSAGRRACLGELPLARMELFLFTSLQHFQSFVPTGQPRSHGV 480
 QY 430 FAFLVTPSPYELCAVPR 446
 DB 481 FAFLVTPSPYELCAVPR 497

RESULT 7
 ADB25831
 ID ADB25831 standard; protein; 497 AA.
 XX
 AC ADB25831;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human CYP2D6-related protein #1.
 XX
 KW human; mutant CYP2D6 gene; drug analysis; drug testing.
 XX
 OS Homo sapiens.
 XX
 PN WO2003050282-A1.
 PD 19-JUN-2003.
 XX
 PF 05-DEC-2002; 2002WO-JP012748.
 XX
 PR 06-DEC-2001; 2001JP-00372548.
 XX
 PA (TSUR) TSUMURA & CO.
 XX
 PI Taniyama M, Ogawa K, Tsuchiya N, Hibino T;
 XX
 DR WPI; 2003-505401/47.
 DR N-PSDB; ADB25776.
 XX
 PT Genetic polymorphisms of CYP2D6 gene in human population for analysis of
 PT drug effect on individual patients and testing of new drugs.
 XX
 PS Claim 8; Page 39-42; 75pp; Japanese.
 XX
 CC The invention comprises mutant forms of the human CYP2D6 gene, containing
 CC one or more of the following mutations G125A, C1858T, T2894C and C2875T.
 CC The mutant human CYP2D6 genes of the invention are useful for analysing
 CC the effect of drugs on individual patients and testing of new drugs. The
 CC present amino acid sequence represents a human protein of the invention.
 XX
 SQ Sequence 497 AA;

Query Match 98.1%; Score 2286.5; DB 6; Length 497;
 Best Local Similarity 89.5%; Pred. No. 3.3e-219;
 Matches 445; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

QY 1 MGLEALVPLAVIVAIFLLVLDLMHRRQWRWAARYSGPLPLGGLNLLHVDFTQNTYCFDQ 60
 DB 1 MGLEALVPLAVIVAIFLLVLDLMHRRQWRWAARYSGPLPLGGLNLLHVDFTQNTYCFDQ 60
 QY 61 LRRFGDVFSLOLAWTPVWVNLGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQ--- 117
 DB 61 LRRFGDVFSLOLAWTPVWVNLGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQGVF 120
 QY 118 -----GRPPRPNGLLDDK 129

Db 121 LARYGPWREQRFRFVSTLRNLGLKKSLEQWVTEEAACLCFAAFANHSGRPFPRPNGLLKD 180
 Qy 130 AVSNVIASLTGRRFEYDDPRFLRLDLAQEGKKEESGFLREVLNAVPLVLIHIPALAGKV 189
 Db 181 AVSNVIASLTGRRFEYDDPRFLRLDLAQEGKKEESGFLREVLNAVPLVLIHIPALAGKV 240
 Qy 190 LRFQKAFITQLDELTEHRTWDPAPPRDLTEAFLAEMERAKGNPSSFNDENLRIYVA 249
 Db 241 LRFQKAFITQLDELTEHRTWDPAPPRDLTEAFLAEMERAKGNPSSFNDENLRIYVA 300
 Qy 250 DLFSAGMVTSTTLAWGLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 309
 Db 301 DLFSAGMVTSTTLAWGLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 360
 Qy 310 HEVORFGDIVPLGVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEKPFRRHPEHF 369
 Db 361 HEVORFGDIVPLGVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEKPFRRHPEHF 420
 Qy 370 LDAQGHFVKPEAFLPFSAGRRACLGELPLARMELFLFTSLLOHFSFSVPTGQPRSHHG 429
 Db 421 LDAQGHFVKPEAFLPFSAGRRACLGELPLARMELFLFTSLLOHFSFSVPTGQPRSHHG 480
 Qy 430 FAFLVTPSPYELCAVPR 446
 Db 481 FAFLVTPSPYELCAVPR 497

RESULT 8
 AAR72377
 ID AAR72377 standard; protein; 497 AA.
 XX
 AC AAR72377;
 XX
 DT 25-MAR-2003 (revised)
 DT 15-NOV-1995 (first entry)
 XX
 DE Human auxiliary cytochrome P450 species 2D6 variant 2 protein.
 XX
 KW Human cytochrome P450; amplification; PCR; primer; expression vector;
 KW yeast NADPH-P450 reductase; safety; fusion protein; metabolite;
 KW carcinogen; mutagen; liver metabolism.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 296 /note= "Cys to Arg variation"
 FT Misc-difference 486 /note= "Thr to Ser variation"
 FT
 XX
 PN EP644267-A2.
 XX
 XX
 PD 22-MAR-1995.
 XX
 XX 20-JUL-1994; 94EP-00111298.
 XX
 PR 20-JUL-1993; 93JP-00201120.
 PR 21-JUL-1993; 93JP-00180246.
 PR 30-JUL-1993; 93JP-00208279.
 XX
 XX (SUMO) SUMITOMO CHEM CO LTD.
 XX
 PI Hayashi K, Sakaki T, Yabusaki Y, Komai K, Kaneko H, Nakatsuka I;
 XX
 DR WPI; 1995-116991/16.
 DR N-PSDB; AAQ87731.
 XX
 PT Evaluation of safety of a chemical cpd. - using recombinant yeast
 PT expressing human cytochrome P450 and a yeast NADPH-P450 reductase.
 XX
 PS Example; Page 91-93; 124pp; English.
 XX
 CC The amino acid sequence of the human auxiliary cytochrome P450 species

CC 2D6 variant 2. This variant contains variations at residues 296: Cys to Arg and 486: Thr to Ser, caused by variations at bases 886: T to C and 1457: C to G in the DNA sequence. The cDNA was amplified by PCR using the primers AAQ87763-6. The product was cloned into the yeast expression vectors pAAHNS or pAHR to produce the vectors p2D6 variant 2 for co-expression of the cytochrome P450 alone or p2D6R variant 2 for co-expression with the yeast NADPH-P450 reductase. The vectors are used in a method for evaluating the safety of a chemical compound by reacting the chemical compound with recombinantly produced human cytochrome P450 molecular species 1A2 (AAQ87714), 2C9 (AAQ87715), 2E1 (AAQ87716), or 3A4 (AAQ87718), or their auxiliary species and variants (AAQ87718-32), and yeast NADPH-P450 reductase, either as a fused protein or in cell extracts, and analyzing the resulting metabolite to assess the safety of the chemical compound. The method is useful for determining whether the chemical compound, or its metabolite, will be converted into a carcinogenic or mutagenic form through metabolism in the liver. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 497 AA;
 Query Match 98.1%; Score 2285.5; DB 2; Length 497;
 Best Local Similarity 89.3%; Pred. No. 4.2e-219;
 Matches 444; Conservative 1; Mismatches 1; Indels 51; Gaps 1;
 Qy 1 MGLEALVPLAVIVAI FLLLLVDLMHRRQRWAARYSPGLPLPGLGNLLHVDVQNTPYCFDQ 60
 Db 1 MGLEALVPLAVIVAI FLLLLVDLMHRRQRWAARYSPGLPLPGLGNLLHVDVQNTPYCFDQ 60
 Qy 61 LRRFGDVFSIQLAWTPVVLNGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQGVF 117
 Db 61 LRRFGDVFSIQLAWTPVVLNGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQGVF 120
 Qy 118 -----GRFPRNGLLDK 129
 Db 121 LARYGPWREQRFRFVSTLRNLGLKKSLEQWVTEEAACLCFAAFANHSGRPFPRPNGLLKD 180
 Qy 130 AVSNVIASLTGRRFEYDDPRFLRLDLAQEGKKEESGFLREVLNAVPLVLIHIPALAGKV 189
 Db 181 AVSNVIASLTGRRFEYDDPRFLRLDLAQEGKKEESGFLREVLNAVPLVLIHIPALAGKV 240
 Qy 190 LRFQKAFITQLDELTEHRTWDPAPPRDLTEAFLAEMERAKGNPSSFNDENLRIYVA 249
 Db 241 LRFQKAFITQLDELTEHRTWDPAPPRDLTEAFLAEMERAKGNPSSFNDENLRIYVA 300
 Qy 250 DLFSAGMVTSTTLAWGLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 309
 Db 301 DLFSAGMVTSTTLAWGLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 360
 Qy 310 HEVORFGDIVPLGVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEKPFRRHPEHF 369
 Db 361 HEVORFGDIVPLGVTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEKPFRRHPEHF 420
 Qy 370 LDAQGHFVKPEAFLPFSAGRRACLGELPLARMELFLFTSLLOHFSFSVPTGQPRSHHG 429
 Db 421 LDAQGHFVKPEAFLPFSAGRRACLGELPLARMELFLFTSLLOHFSFSVPTGQPRSHHG 480
 Qy 430 FAFLVTPSPYELCAVPR 446
 Db 481 FAFLVTPSPYELCAVPR 497

RESULT 9
 AAR93184
 ID AAR93184 standard; protein; 497 AA.
 XX
 AC AAR93184;
 XX
 DT 11-OCT-1996 (first entry)
 XX
 DE Human cytochrome P450 molecular species 2D6 variant #3 protein.
 KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;
 KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;

Db 121 LARYGPAMRQRRFSVSTLRNLGLGKSLQWVTEAAACLCAPFANHSRPRPNGLLDK 180
Qy 130 AVSNVIASLTGRRREYDDPRFLRLDLAQEGLKEESGFLREVLNAVPLVLIHPALAGKV 189
Db 181 AVSNVIASLTGRRREYDDPRFLRLDLAQEGLKEESGFLREVLNAVPLVLIHPALAGKV 240
Qy 190 LRFOKAFITQDLDELTEHRMTWDPAPPRDLTEAFLAEMKAKGNPSSFNENLRIVVA 249
Db 241 LRFOKAFITQDLDELTEHRMTWDPAPPRDLTEAFLAEMKAKGNPSSFNENLRIVVA 300
Qy 250 DLFSAGMVTSTTLAWGLLMILHDPVQRRVQOEIDVIGVRRPEMGDQAHMPYTTAVI 309
Db 301 DLFSAGMVTSTTLAWGLLMILHDPVQRRVQOEIDVIGVRRPEMGDQAHMPYTTAVI 360
Qy 310 HEVQRFQDIVPLGVTHMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPRFRPHPHF 369
Db 361 HEVQRFQDIVPLGVTHMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPRFRPHPHF 420
Qy 370 LDAQGHFVKPEAFPLFSAGRRACLGELPLARWELFLFTSLLOHFSFSVPTGQPRPSHGV 429
Db 421 LDAQGHFVKPEAFPLFSAGRRACLGELPLARWELFLFTSLLOHFSFSVPTGQPRPSHGV 480
Qy 430 FAFLVTPSPYELCAVPR 446
Db 481 FAFLVTPSPYELCAVPR 497

RESULT 11
ID ABB09563
XX ABB09563 standard; protein; 497 AA.
AC ABB09563;
XX
XX
XX 02-SEP-2002 (first entry)
XX Human CYP2D6 protein, SEQ ID NO:3.
DE
DE Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;
KW chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;
KW antiarrhythmic; arrhythmia; adrenoceptor antagonist; hypertension;
KW tricyclic antidepressant; procainamide; drug induced lupus syndrome;
KW environmentally linked disease; Parkinson's disease; haplotyping;
KW genotyping; haplotype; genetic variant; single nucleotide polymorphism;
KW SNP; drug screening; drug discovery.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FT Misc-difference 7
FT /note= "Encoded by RTG in the sequence given in ABO72215.
FT This residue is Met (encoded by ATG) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT
FT Misc-difference 11
FT /note= "Encoded by RTG in the sequence given in ABO72215.
FT This residue is Met (encoded by ATG) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT
FT Misc-difference 34
FT /note= "Encoded by YCA in the sequence given in ABO72215.
FT This residue is Ser (encoded by TCA) rather than Pro
FT (encoded by CCA) in a polymorphic variant"
FT
FT Misc-difference 88
FT /note= "Encoded by CRC in the sequence given in ABO72215.
FT This residue is His (encoded by CAC) rather than Arg
FT (encoded by CGC) in a polymorphic variant"
FT
FT Misc-difference 91
FT /note= "Encoded by MTG in the sequence given in ABO72215.
FT This residue is Met (encoded by ATG) rather than Leu
FT (encoded by CTG) in a polymorphic variant"
FT
FT Misc-difference 94
FT /note= "Encoded by CRC in the sequence given in ABO72215.
FT This residue is Arg (encoded by CGC) rather than His
FT (encoded by CAC) in a polymorphic variant"
FT
FT Misc-difference 98
FT

FT /note= "Encoded by ACS in the sequence given in ABO72215"
FT Misc-difference 104
FT /note= "Encoded by GYG in the sequence given in ABO72215.
FT This residue is Ala (encoded by GCG) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT
FT Misc-difference 107
FT /note= "Encoded by WYC in the sequence given in ABO72215.
FT This residue is Phe (encoded by TTC) rather than Thr
FT (encoded by ACC) in a polymorphic variant"
FT
FT Misc-difference 109
FT /note= "Encoded by RTC in the sequence given in ABO72215.
FT This residue is Val (encoded by GTC) rather than Ile
FT (encoded by ATC) in a polymorphic variant"
FT
FT Misc-difference 111
FT /note= "Encoded by GGY in the sequence given in ABO72215"
FT
FT Misc-difference 112
FT /note= "Encoded by TTY in the sequence given in ABO72215"
FT
FT Misc-difference 120
FT /note= "Encoded by WTC in the sequence given in ABO72215.
FT This residue is Ile (encoded by ATC) rather than Phe
FT (encoded by TTC) in a polymorphic variant"
FT
FT Misc-difference 128
FT /note= "Encoded by YGG in the sequence given in ABO72215.
FT This residue is Arg (encoded by CGG) rather than Trp
FT (encoded by TGG) in a polymorphic variant"
FT
FT Misc-difference 136
FT /note= "Encoded by RTS in the sequence given in ABO72215.
FT This residue is Ile (encoded by ATC) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT
FT Misc-difference 151
FT /note= "Encoded by SAG in the sequence given in ABO72215.
FT This residue is Glu (encoded by GAG) rather than Gln
FT (encoded by CAG) in a polymorphic variant"
FT
FT Misc-difference 155
FT /note= "Encoded by RAG in the sequence given in ABO72215.
FT This residue is Lys (encoded by AAG) rather than Glu
FT (encoded by GAG) in a polymorphic variant"
FT
FT Misc-difference 232
FT /note= "Encoded by CAY in the sequence given in ABO72215"
FT
FT Misc-difference 338
FT /note= "Encoded by RTG in the sequence given in ABO72215.
FT This residue is Met (encoded by ATG) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT
FT Misc-difference 344
FT /note= "Encoded by YCA in the sequence given in ABO72215.
FT This is a stop codon (TGA) rather than Arg (encoded by
FT CGA) in a polymorphic variant"
FT
FT Misc-difference 361
FT /note= "Encoded by CAY in the sequence given in ABO72215"
FT
FT Misc-difference 486
FT /note= "Encoded by ASC in the sequence given in ABO72215.
FT This residue is Thr (encoded by ACC) rather than Ser
FT (encoded by AGC) in a polymorphic variant"
FT
FT Misc-difference 497
FT /note= "Encoded by YGC in the sequence given in ABO72215.
FT This residue is Cys (encoded by TGC) rather than Arg
FT (encoded by CGC) in a polymorphic variant"
FT
XX WO200238589-A2.
XX 16-MAY-2002.
XX
XX 09-NOV-2001; 2001WO-US047396.
XX
XX 09-NOV-2000; 2000US-0247943P.
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Anastasio AE, Chew A, Choi JY, Denton RR, Nandabalan K;
XX Petersen N, Rounds E;
XX WPI; 2002-519292/55.
XX N-PSDB; ABO72215, ABO72216, ABO72364.
DR

XX Novel genetic variants of Cytochrome P450, Subfamily IID, Polypeptide 6
PT isogenes, useful for improving efficiency and reliability in drug
PT development for treating hypertension, arrhythmias and Parkinson's
PT disease.

XX Claim 29; Fig 3; 158pp; English.

CC The invention relates to a method for haplotyping the cytochrome P450,
CC subfamily IID, polypeptide 6 (CYP2D6) gene (AB072215, AB072364) of an
CC individual, and also describes 29 novel polymorphic sites within the
CC human CYP2D6 gene. The CYP2D6 gene is located on chromosome 22q13.1 and
CC contains 9 exons which encode a 497 amino acid protein (AB09563). CYP2D6
CC is a mono-oxygenase involved in the detoxification of many drugs and
CC environmental chemicals. It plays a role in the metabolism of drugs such
CC as antiarrhythmics, adrenoceptor antagonists and tricyclic
CC antidepressants, and is also involved in the formation of a metabolite
CC linked to the drug-induced lupus syndrome observed with procainamide.
CC Variations in CYP2D6 activity or expression may also influence an
CC individual's susceptibility to environmentally-linked diseases, and it
CC has been demonstrated that CYP2D6 activity may be involved in the
CC pathogenesis of Parkinson's disease, with individuals with a less active
CC form of the enzyme tending to have an earlier onset of this condition.
CC CYP2D6 nucleic acid sequences are useful in studying the expression and
CC function of CYP2D6, and in expressing CYP2D6 protein for use in screening
CC drugs for the treatment of CYP2D6-associated diseases (e.g.,
CC hypertension, atrial and ventricular arrhythmias, Parkinson's disease,
CC and drug-induced lupus syndrome) or which are metabolised by CYP2D6.
CC CYP2D6 nucleic acids and proteins are also useful in studying the effect
CC of polymorphisms on the biological activity of CYP2D6. Polymorphisms in
CC the target region may be determined by the use of allele-specific
CC oligonucleotides (ASOs; AB072217-AB072303) as probes and primers, and by
CC primer extension using oligonucleotide primers comprising sequences
CC AB072304-AB072361. The method of the invention is useful for haplotyping
CC the CYP2D6 gene in populations and in individuals, enabling decisions to
CC be made as to whether CYP2D6 is a likely therapeutic target for a disease
CC of interest, and to control for genetically-based bias in the design of
CC drugs that target or are metabolised by CYP2D6. In addition, transgenic
CC animals comprising a human CYP2D6 gene are useful for studying the
CC expression of CYP2D6 isogenes in vivo, for in vivo screening and testing
CC of drugs targeted to or metabolised by CYP2D6, and for testing the
CC efficacy of therapeutic agents and compounds for treating CYP2D6-
CC associated conditions in a biological system. The present sequence
CC represents the specifically claimed human CYP2D6 protein. This sequence
CC contains 18 polymorphic sites caused by polymorphisms in the coding
CC sequence (AB072216)

XX Sequence 497 AA;

Query Match 98.1%; Score 2285.5; DB 5; Length 497;
Best Local Similarity 89.3%; Pred. No. 4.2e-219;
Matches 444; Conservative 1; Mismatches 1; Indels 51; Gaps 1;

QY 1 MGLEALVPLAVIVAIFLLVLDLMHRRQWAAARYSPGFLPLGLNLLHVDVFNTPYCFDQ 60
DB 1 MGLEALVPLAVIVAIFLLVLDLMHRRQWAAARYSPGFLPLGLNLLHVDVFNTPYCFDQ 60
QY 61 LRRRFGDVFSQAWTPVVVNLGLAAVREALVTHGEDTADRPVPTQILGFGPRSQ--- 117
DB 61 LRRRFGDVFSQAWTPVVVNLGLAAVREALVTHGEDTADRPVPTQILGFGPRSQGVF 120
QY 118 -----GRPRFNPGLLDK 129
DB 121 LARYGPWREQRFSVSTLRNLGLKKSLEQWVTEAACICAAFNHSGRPRFNPGLLDK 180
QY 130 AVSNVITASLTCGRFFYDDPRFLRLDLAQEGLKEESGFLREVNLNVPVLLHIALAGKV 189
DB 181 AVSNVITASLTCGRFFYDDPRFLRLDLAQEGLKEESGFLREVNLNVPVLLHIALAGKV 240
QY 190 LRFQKAFITQDDELLTEHRTWDPQPPRDLTEAFIAEMEAKGNPSSFNENRIIYVA 249
DB 241 LRFQKAFITQDDELLTEHRTWDPQPPRDLTEAFIAEMEAKGNPSSFNENRIIYVA 300

QY 250 DLFSAGMVTSTTTLAWGLLMLHPDVQRRVQOEIDDDVIGQVRRPMPGQAHMPYTTAVI 309
DB 301 DLFSAGMVTSTTTLAWGLLMLHPDVQRRVQOEIDDDVIGQVRRPMPGQAHMPYTTAVI 360
QY 310 HEVQRFGDIVPLGVTHMTSRDIEVQGFRIIPKGTITLNLSSVLKDEAVWEKPRFRPHEHF 369
DB 361 HEVQRFGDIVPLGVTHMTSRDIEVQGFRIIPKGTITLNLSSVLKDEAVWEKPRFRPHEHF 420
QY 370 LDAQGHFVKPEAPLFPFSAGRRACLGELPRLARMELFFFTSLLOHFSFVSPTGPRPSHHGV 429
DB 421 LDAQGHFVKPEAPLFPFSAGRRACLGELPRLARMELFFFTSLLOHFSFVSPTGPRPSHHGV 480
QY 430 FAPLVTPSPSYELCAVPR 446
DB 481 FAPLVSPSYELCAVPR 497

RESULT 12

AA026405
ID AA026405 standard; protein; 497 AA.

AC AA026405;

DT 30-JAN-2003 (first entry)

DE Human drug-metabolising enzyme related protein.

KW Cytostatic; antiparkinsonian; gene therapy; cancer; Parkinson's disease;
KW human drug-metabolising protein; enzyme.

OS Homo sapiens.

PN WO2002792333-A1.

PD 10-OCT-2002.

PF 01-APR-2002; 2002WO-US009738.

PR 30-MAR-2001; 2001US-00820788.

PA (PEKE) PE CORP NY.

PA (DFRA/) DI FRANCESCO V.

PA (BEAS/) BEASLEY E M.

PI Shao W, Yan C;

DR WPI; 2003-040649/03.

PT New human drug-metabolizing proteins and nucleic acids related to the
PT Cytochrome P450 IID drug-metabolizing enzyme subfamily, useful for
PT treating a condition mediated by a human enzyme protein e.g., cancer.

PS Disclosure; Fig 2B; 72pp; English.

CC The invention relates to a novel isolated polypeptide comprising a 446-
CC amino acid sequence or its allelic variant, orthologue or fragment. The
CC allelic variant or orthologue is encoded by a nucleic acid that
CC hybridises under stringent conditions to the opposite strand of the
CC nucleic acid comprising a sequence having 1537 or 10278 bp. The fragment
CC comprises at least 10 contiguous amino acids of the 446-amino acid
CC sequence. The polypeptide is useful for preparing a pharmaceutical
CC composition for treating a disease or condition mediated by a human
CC enzyme protein, e.g. cancer or Parkinson's disease. This sequence
CC represents a human drug-metabolising related protein of the invention

XX Sequence 497 AA;

Query Match 98.1%; Score 2285.5; DB 6; Length 497;

Best Local Similarity 89.3%; Pred. No. 4.2e-219;

Matches 444; Conservative 1; Mismatches 1; Indels 51; Gaps 1;

QY 1 MGLEALVPLAVIVAIFLLVLDLMHRRQWAAARYSPGFLPLGLNLLHVDVFNTPYCFDQ 60
|||||

Db 1 MGLEALVPLAVIAVIFLLVLDLMHRRQWAARYPPGCLPLPGLGNLLHVDFTQNTPCFDQ 60
QY 61 LRRRFGDVFSLOLAWTPVVLNGLAAVREALVTHGEDTADRPVPIITQILGFGPRSQ--- 117
Db 61 LRRRFGDVFSLOLAWTPVVLNGLAAVREALVTHGEDTADRPVPIITQILGFGPRSQGVF 120
QY 118 -----GRPRPNGLLDK 129
Db 121 LARYGPAMREORRFSVSTLRNLGLGKKSLEQWVTEEAACLCAPAFANHSGRPRFNGLLDK 180
QY 130 AVSNVIASLTGRRFEYDDPRFLRLDLAQEGKESGFLREVNAVPLVLLHI PALAGKV 189
Db 181 AVSNVIASLTGRRFEYDDPRFLRLDLAQEGKESGFLREVNAVPLVLLHI PALAGKV 240
QY 190 LRFQKAFLTQDLDELTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPESFNENLRIYVA 249
Db 241 LRFQKAFLTQDLDELTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPESFNENLRIYVA 300
QY 250 DLFSAGMVTSTTLAWGLLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 309
Db 301 DLFSAGMVTSTTLAWGLLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 360
QY 310 HEVQRFQGDIVPLGVTHMTSRDIEVOGFRI PKGTTLITNLSSVLKDEAVWEKPFPRFPHSHF 369
Db 361 HEVQRFQGDIVPLGVTHMTSRDIEVOGFRI PKGTTLITNLSSVLKDEAVWEKPFPRFPHSHF 420
QY 370 LDAQGHFVKPEAFLPFSAGRRACI GEPLARMELFLFTSLLOHFSFSVPTGQPRSHHG 429
Db 421 LDAQGHFVKPEAFLPFSAGRRACI GEPLARMELFLFTSLLOHFSFSVPTGQPRSHHG 480
QY 430 FAFVTPSPYELCAVPR 446
Db 481 FAFVTPSPYELCAVPR 497

RESULT 13

ADB25833
ID ADB25833 standard; protein; 497 AA.

XX AC ADB25833;

XX DT 20-NOV-2003 (first entry)

XX DE Human CYP2D6-related protein #3.

XX KW human; mutant CYP2D6 gene; drug analysis; drug testing.

XX OS Homo sapiens.

XX PN WO2003050282-A1.

XX PD 19-JUN-2003.

XX PF 05-DEC-2002; 2002WO-JP012748.

XX PR 06-DEC-2001; 2001JP-00372548.

XX XX (TSUR) TSUMURA & CO.

XX PI Taniyama M, Ogawa K, Tsuchiya N, Hibino T;

XX DR WPI: 2003-505401/47.

XX DR N-PSDB; ADB25778.

XX PT Genetic polymorphisms of CYP2D6 gene in human population for analysis of

XX PT drug effect on individual patients and testing of new drugs.

XX PS Claim 8; Page 46-50; 75pp; Japanese.

XX CC The invention comprises mutant forms of the human CYP2D6 gene, containing
CC one or more of the following mutations G125A, C1858T, T2874C and C2875T.
CC The mutant human CYP2D6 genes of the invention are useful for analysing
CC the effect of drugs on individual patients and testing of new drugs. The

CC present amino acid sequence represents a human protein of the invention.
XX SQ Sequence 497 AA;

Query Match 98.1%; Score 2285.5; DB 6; Length 497;
Best Local Similarity 89.3%; Pred. No. 4.2e-219;
Matches 444; Conservative 1; Mismatches 1; Indels 51; Gaps 1;

QY 1 MGLEALVPLAVIAVIFLLVLDLMHRRQWAARYSPGCLPLPGLGNLLHVDFTQNTPCFDQ 60
Db 1 MGLEALVPLAVIAVIFLLVLDLMHRRQWAARYSPGCLPLPGLGNLLHVDFTQNTPCFDQ 60
QY 61 LRRRFGDVFSLOLAWTPVVLNGLAAVREALVTHGEDTADRPVPIITQILGFGPRSQ--- 117
Db 61 LRRRFGDVFSLOLAWTPVVLNGLAAVREALVTHGEDTADRPVPIITQILGFGPRSQGVF 120
QY 118 -----GRPRPNGLLDK 129
Db 121 LARYGPAMREORRFSVSTLRNLGLGKKSLEQWVTEEAACLCAPAFANHSGRPRFNGLLDK 180
QY 130 AVSNVIASLTGRRFEYDDPRFLRLDLAQEGKESGFLREVNAVPLVLLHI PALAGKV 189
Db 181 AVSNVIASLTGRRFEYDDPRFLRLDLAQEGKESGFLREVNAVPLVLLHI PALAGKV 240
QY 190 LRFQKAFLTQDLDELTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPESFNENLRIYVA 249
Db 241 LRFQKAFLTQDLDELTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPESFNENLRIYVA 300
QY 250 DLFSAGMVTSTTLAWGLLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 309
Db 301 DLFSAGMVTSTTLAWGLLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 360
QY 310 HEVQRFQGDIVPLGVTHMTSRDIEVOGFRI PKGTTLITNLSSVLKDEAVWEKPFPRFPHSHF 369
Db 361 HEVQRFQGDIVPLGVTHMTSRDIEVOGFRI PKGTTLITNLSSVLKDEAVWEKPFPRFPHSHF 420
QY 370 LDAQGHFVKPEAFLPFSAGRRACI GEPLARMELFLFTSLLOHFSFSVPTGQPRSHHG 429
Db 421 LDAQGHFVKPEAFLPFSAGRRACI GEPLARMELFLFTSLLOHFSFSVPTGQPRSHHG 480
QY 430 FAFVTPSPYELCAVPR 446
Db 481 FAFVTPSPYELCAVPR 497

RESULT 14

ABU09593
ID ABU09593 standard; protein; 497 AA.

XX AC ABU09593;

XX DT 16-JUL-2003 (first entry)

XX DE Human cytochrome p450 gene CYP2D6, wild-type protein.

XX KW Human; cytochrome P450; CYP2D6; chromosome 22; SNP; drug metabolism;
KW cardiovascular disorder; psychiatric disorder; drug sensitivity.

XX OS Homo sapiens.

XX PN EP1281755-A2.

XX PD 05-FEB-2003.

XX PF 16-JUL-2002; 2002EP-00254972.

XX PR 31-JUL-2001; 2001US-0309111P.

XX PA (PFIZ) PFIZER PROD INC.

XX PI Milos PM, Webb SM;

XX DR WPI: 2003-373769/36.

DR N-PSDB; ACA61303.
XX New cytochrome P450 2D6 gene variants and polypeptides, useful for
PT determining if a subject has or is at risk of developing a drug
PT sensitivity condition or disorder that is associated with an aberrant
PT CYP2D6 activity.
XX
PS Claim 11; Fig 4; 89pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising a cytochrome
CC P450 2D6 gene variant, e.g. G5799C or C5816AT (referring to the genomic
CC sequence or the same variant nucleotide in the corresponding cDNA
CC sequences). Also included are probes, primers (allele specific
CC oligonucleotides) and arrays used to detect and/or amplify the CYP2D6
CC gene polymorphic regions, the variant polypeptides, antibodies which are
CC capable of distinguishing between the variant and wild-type polypeptides,
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
CC determining whether an individual is susceptible to being a poor
CC metaboliser of drugs. The DNA probe is useful for hybridising to a
CC variant form of the CYP2D6 gene. The primer is useful for amplifying the
CC C5816TA allelic variant. The allele specific nucleotide is useful for the
CC detection of the C5816TA allelic variant. The methods are useful for
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and
CC determining if an individual is susceptible to being a poor metaboliser
CC of drugs. The nucleic acids are useful as probes or primers for
CC determining whether a subject has a genetic deficiency for metabolising
CC drugs that are substrates of P450 CYP2D6. The methods are useful for
CC determining if a subject has or is at risk of developing a drug
CC sensitivity condition or disorder that is associated with an aberrant
CC CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
CC aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
CC appropriate drugs or determining the course of treatment to administer to
CC a subject to treat cardiovascular or psychiatric disorders, or for
CC treating a subject with a drug sensitivity or disorder associated with a
CC specific allelic variant of a polymorphic region of the CYP2D6 gene. The
CC antibodies are useful for monitoring CYP2D6 protein levels in an
CC individual for determining whether a subject has a disease or conditions
CC associated with an aberrant CYP2D6 protein level. The gene is located on
CC human chromosome 22. The present sequence is the wild-type CYP2D6 protein
XX
SQ Sequence 497 AA;

Query Match 98.0%; Score 2282.5; DB 6; Length 497;
Best Local Similarity 89.1%; Pred. No. 8.4e-219;
Matches 443; Conservative 2; Mismatches 1; Indels 51; Gaps 1;

QY 1 MGLEALVPLAVIIVAI FLLLVLMHRRQRWAARYSPGFLPLGLGNLLHVDYDQNTPYCFDQ 60
DB 1 MGLEALVPLAVIIVAI FLLLVLMHRRQRWAARYSPGFLPLGLGNLLHVDYDQNTPYCFDQ 60
QY 61 LRRFGDVSLOLAWTPVVVINGLAAREALVTHGSDTADRPVPTIILGGRSQ--- 117
DB 61 LRRFGDVSLOLAWTPVVVINGLAAREALVTHGSDTADRPVPTIILGGRSQGVF 120
QY 118 -----GRPRPNGLLDK 129
DB 121 LARYGPANREQRFSVSTLNLGLGKSLSEQWTEAAGLCAAFANHSRPRPNGLLDK 180
QY 130 AVSNVIASLTGRRFEYDPRFLRLDLAQEGKESGFLREVLNVAVPVLLHIPALAGKV 189
DB 181 AVSNVIASLTGRRFEYDPRFLRLDLAQEGKESGFLREVLNVAVPVLLHIPALAGKV 240
QY 190 LRFQKAFITOLDLLETHRTWDPAPPRDLTAFLAEMKACNPFSSFNENLRIVVA 249
DB 241 LRFQKAFITOLDLLETHRTWDPAPPRDLTAFLAEMKACNPFSSFNENLRIVVA 300
QY 250 DLFSAGWVTYTTTLWGLLMLHPDVORRVQCEIDVIGVRRRPMGQAHMPTTAVI 309
DB 301 DLFSAGWVTYTTTLWGLLMLHPDVORRVQCEIDVIGVRRRPMGQAHMPTTAVI 360
QY 310 HEVQRFEDI VPLGVTMTSRDIEVOGFRI PKGTTLITNLSSVLKDEAVWEKPFRRFHPHF 369

DB 361 HEVQRFEDI VPLGVTMTSRDIEVOGFRI PKGTTLITNLSSVLKDEAVWEKPFRRFHPHF 420
QY 370 LDAQGHFVKPEAFLPFSAGRRACLEPLARMEFLFFTSLLQHFSPVPTGPRSHHGV 429
DB 421 LDAQGHFVKPEAFLPFSAGRRACLEPLARMEFLFFTSLLQHFSPVPTGPRSHHGV 480
QY 430 FAFVLTPSPYELCAVPR 446
DB 481 FAFVLTPSPYELCAVPR 497
RESULT 15
ABR82026
ID ABR82026 standard; protein; 497 AA.
XX ABR82026;
AC ABR82026;
XX 22-SEP-2003 (first entry)
DT Human cytochrome P450 2D6 amino acid sequence.
DE Human; protein array; protein moiety; phenotype; drug discovery;
DE naturally occurring variant; pharmacogenomic; diagnostic;
KW parallel analysis; tumour suppressor; p53; cytochrome P450.
KW Homo sapiens.
OS
XX WO2003048768-A2.
PN 12-JUN-2003.
PD
XX 05-DEC-2002; 2002WO-GB005499.
PF
XX 05-DEC-2001; 2001US-0335806P.
PR 16-SEP-2002; 2002US-0410815P.
XX (SENS-) SENSE PROTEOMIC LTD.
XX Boutell JM, Godber BLJ, Hart DJ, Blackburn JD;
XX WPI; 2003-569063/53.
DR N-PSDB; ACF06056.
XX
PT New protein array, useful for determining the phenotype of a naturally
PT occurring variant of a DNA sequence of interest, comprises a surface upon
PT which at least two protein moieties are deposited.
XX
PS Example 5; Fig 13B; 84pp; English.
XX
CC The present invention describes a protein array comprising a surface upon
CC which at least two protein moieties are deposited at spatially defined
CC locations, where the protein moieties are naturally occurring variants of
CC a DNA sequence of interest. Also described: (1) making a protein array;
CC (2) screening a set of protein moieties for molecules that interact with
CC one or more proteins; and (3) simultaneously determining the relative
CC properties of members of a set of protein moieties. The protein array can
CC be used for determining the phenotype of a naturally occurring variant of
CC a DNA sequence of interest. The protein array is useful for drug
CC discovery, pharmacogenomics and diagnostics. The protein array allows the
CC parallel analysis of closely related proteins with a sensitivity that is
CC at least comparable to existing methods, if not better, with small
CC volumes of potentially expensive ligands, and in a quantitative,
CC comparative functional analysis manner not previously possible. ACP06000
CC to ACF06056 and ABR81975 to ABR82026 represent sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 497 AA;

Query Match 98.0%; Score 2282.5; DB 6; Length 497;
Best Local Similarity 89.1%; Pred. No. 8.4e-219;
Matches 443; Conservative 2; Mismatches 1; Indels 51; Gaps 1;

Qy	1	MGLEALVPLAVIVAIFLLLLVDLMHRRORWAARYSPGPLPLGLGNLLHVDVFQNTYCFDQ	60
Db	1	MGLEALVPLAVIVAIFLLLLVDLMHRRORWAARYPPGPLPLGLGNLLHVDVFQNTYCFDQ	60
Qy	61	LRRRFGDVFSQLAWTPVVLINGLAAREALVTHGEDTADRPVPITQILGFGPRSQ---	117
Db	61	LRRRFGDVFSQLAWTPVVLINGLAAREALVTHGEDTADRPVPITQILGFGPRSQGVF	120
Qy	118	-----GRPRRPNGLLDDK	129
Db	121	LARYGPAREQRRSVSTLRNLGLGKKSLEQWTEBSAACLAFAFNSGRPRPNGLLDDK	180
Qy	130	AVSNVIAISLTCGRFEYDDPRFLRLDLLAQEGLEKESGFLREVINAVPVLLHIIPALAGKV	189
Db	181	AVSNVIAISLTCGRFEYDDPRFLRLDLLAQEGLEKESGFLREVINAVPVLLHIIPALAGKV	240
Qy	190	LRFKQAFITQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEKAKGNPESFNENDLRIVVA	249
Db	241	LRFKQAFITQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEKAKGNPESFNENDLRIVVA	300
Qy	250	DLFSAGMWTSTTTLAWGLLLMLHPDVORRVQOEIDDVIGQVRRPBMGQOAMPPYTAVI	309
Db	301	DLFSAGMWTSTTTLAWGLLLMLHPDVORRVQOEIDDVIGQVRRPBMGQOAMPPYTAVI	360
Qy	310	HEVQRFQGDIVPLGVTHMTSRDIEVQGRFRIPKGTTTLITNLSSVLKDEAVWEKFPFRHPHF	369
Db	361	HEVQRFQGDIVPLGVTHMTSRDIEVQGRFRIPKGTTTLITNLSSVLKDEAVWEKFPFRHPHF	420
Qy	370	LDAQGHFVKPBEAFLPFSAGRRACLGEPALAMELFLFTTSLLQHFVSFVTPQGRPSHHGV	429
Db	421	LDAQGHFVKPBEAFLPFSAGRRACLGEPALAMELFLFTTSLLQHFVSFVTPQGRPSHHGV	480
Qy	430	FAFLVTPSPSYELCAVPR	446
Db	481	FAFLVTPSPSYELCAVPR	497

Search completed: February 25, 2004, 04:29:32
Job time : 64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 04:11:47 ; Search time 21 Seconds
(without alignments)
2042.925 Million cell updates/sec

Title: US-09-820-788A-2
Perfect score: 2330
Sequence: 1 MGLEALVPLAVIAIFLLLV.....HGVPFLVTPSPYELCAVPR 446

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: Piri: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2282.5	98.0	497	04HUD1	debrisoquine 4-hyd
2	2140.5	91.9	497	G02938	probable debrisoqu
3	1769	75.9	500	S37284	cytochrome P450 2D
4	1751	75.2	500	J02558	cytochrome P450 2D
5	1739	74.6	500	1 JC5819	cytochrome P450 2D
6	1724.5	74.0	500	D31579	cytochrome P450 2D
7	1697	72.8	500	JE0259	cytochrome P450 2D
8	1690	72.5	500	1 JC4157	cytochrome P450 2D
9	1638.5	70.3	500	S16872	cytochrome P450 2D
10	1625	69.7	500	B26822	cytochrome P450 2D
11	1614.5	69.3	504	04RTD5	cytochrome P450 2D
12	1614.5	69.3	504	A26822	debrisoquine 4-hyd
13	1611	69.1	500	1 JC4153	cytochrome P450 2D
14	1597.5	68.6	504	A27384	steroid 16alpha-hy
15	1580.5	67.8	504	I49427	cytochrome P450 16
16	1572.5	67.5	504	A30247	cytochrome P450 16
17	1567.5	67.3	504	1 I49428	cytochrome P450 16
18	1456	62.5	505	S19169	cytochrome P450 2B
19	821.5	35.3	491	I49625	testosterone 16a-h
20	816.5	35.0	491	A27717	cytochrome P450 2B
21	812.5	34.9	491	2 S31277	cytochrome P450 2B
22	812	34.8	491	2 I84735	testosterone 16a-h
23	810.5	34.8	491	04R8PC	cytochrome P450 2B
24	810.5	34.8	491	2 S31278	cytochrome P450 2B
25	808.5	34.7	491	1 04RTPB	cytochrome P450 2B
26	804.5	34.5	501	1 A40938	cytochrome P450 1b
27	801	34.4	494	2 A46588	cytochrome P450 ar
28	800	34.3	491	2 A31047	testosterone 16alp
29	799.5	34.3	491	2 S35666	cytochrome P450 2B

ALIGNMENTS

RESULT 1

O4HUD1
debrisoquine 4-hydroxylase (EC 1.14.14.-) cytochrome P450 2D6 - human
N;Alternate names: CYP2D6; cytochrome P450 isozyme 2D; cytochrome P450db1
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text_change 03-Mar-2000
C;Accession: S01199; A28883; JC4156; A33629; A30335
R;Gonzalez, F.J.; Skoda, R.C.; Kimura, S.; Umeno, M.; Zanger, U.M.; Nebert, D.W.; Gelboin, H.
Nature 331, 442-446, 1988
A;Title: Characterization of the common genetic defect in humans deficient in debrisoqu
A;Reference number: S01199; MUID:88122614; PMID:3123997
A;Accession: S01199
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-497 <CON>
A;Cross-references: EMBL:X08006; NID:G30450; PIDN:CAA30807.1; PID:G30451
R;Gonzalez, F.J.; Vilbois, F.; Hardwick, J.P.; McBride, O.W.; Nebert, D.W.; Gelboin, H.
Genomics 2, 174-179, 1988
A;Title: Human debrisoquine 4-hydroxylase (P450IID1): cDNA and deduced amino acid sequ
A;Reference number: A28883; MUID:88314109; PMID:3410476
A;Accession: A28883
A;Molecule type: mRNA
A;Residues: 1-497 <CON2>
A;Cross-references: EMBL:M20403; NID:G181349; PIDN:AAA52153.1; PID:G181350
R;Jiang, Q.; Voigt, J.M.; Colby, H.D.
Biochem. Biophys. Res. Commun. 209, 1149-1156, 1995
A;Title: Molecular cloning and sequencing of a guinea pig cytochrome P4502D (CYP2D16):
A;Reference number: JC4153; MUID:95251703; PMID:7733969
A;Accession: JC4156
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-497 <JIA>
R;Kimura, S.; Umeno, M.; Skoda, R.C.; Meyer, U.A.; Gonzalez, F.J.
Am. J. Hum. Genet. 45, 889-904, 1989
A;Title: The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification.
A;Reference number: A33629; MUID:90072069; PMID:2574001
A;Accession: A33629
A;Molecule type: DNA
A;Residues: 1-373, 'V', 375-497 <KIM>
A;Cross-references: EMBL:M33388; NID:G181303; PIDN:AAA53500.1; PID:G181304
R;Manns, M.P.; Johnson, E.P.; Griffin, K.J.; Tan, E.M.; Sullivan, K.F.
J. Clin. Invest. 83, 1066-1072, 1989
A;Title: Major antigen of liver kidney microsomal autoantibodies in idiopathic autoimmune
A;Reference number: A30335; MUID:89155788; PMID:2466049
A;Accession: A30335
A;Molecule type: mRNA
A;Residues: 125-373, 'V', 375-485, 'T', 487-497 <MAN>
A;Cross-references: EMBL:M24499; NID:G522194; PIDN:AAA36403.1; PID:G522195
C;Genetics:
A;Gene: GDB:CYP2D6
A;Cross-references: GDB:I32127; OMIM:124030
A;Map position: 22q13.1-22q13.1

unspecific monooxy
cytochrome P450 2B
cytochrome P450 2B
cytochrome P450 11
cytochrome P450 2C
cytochrome P450 -
cytochrome P450 -
cytochrome P450 2C
cytochrome P450 2K
cytochrome P450 2K
cytochrome P450 50
cytochrome P450 2B
cytochrome P450 2B
cytochrome P450 -
cytochrome P450 2B

A:introns: 60/3; 118/1; 169/1; 222/3; 281/3; 329/1; 391/3; 439/1
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F:302-465/Domain: cytochrome P450 homology <CYP>
F:443/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 98.0%; Score 2282.5; DB 1; Length 497;
Best Local Similarity 89.1%; Pred. No. 3.1e-168;
Matches 443; Conservative 2; Mismatches 1; Indels 51; Gaps 1;

QY 1 MGLEALVPLAVIAIFLLLVLMHRRQWAAARYSPGPLPLGGLNLLHVDVFQNTYCFDQ 60
DB 1 MGLEALVPLAVIAIFLLLVLMHRRQWAAARYSPGPLPLGGLNLLHVDVFQNTYCFDQ 60
QY 61 LRRRGDVFSIQAWTPVVLNGLAAVREALVTHGEDTADRPVPIITQILGFGPSQ--- 117
DB 61 LRRRGDVFSIQAWTPVVLNGLAAVREALVTHGEDTADRPVPIITQILGFGPSQGVF 120
QY 118 -----GRPRPNGLLDK 129
DB 121 LARYGPWREORRFSVSTLRNLGLGKKSLEQWVTEAAACLAAPNHSGRPRPNGLLDK 180
QY 130 AVSNVIASITCGRRFEYDDPRFLRLDLAQGLKEESGFLREVLANVPLVLLHIPALAGKV 189
DB 181 AVSNVIASITCGRRFEYDDPRFLRLDLAQGLKEESGFLREVLANVPLVLLHIPALAGKV 240
QY 190 LRFQKAFITQDLDELTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPESFNDENRIVVA 249
DB 241 LRFQKAFITQDLDELTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPESFNDENRIVVA 300
QY 250 DLFSAGMVTSTTLAWGLLMLLHPDVQRRVQOEIDDDVIGVRRPEMGDOAHMPYTTAVI 309
DB 301 DLFSAGMVTSTTLAWGLLMLLHPDVQRRVQOEIDDDVIGVRRPEMGDOAHMPYTTAVI 360
QY 310 HEVQRFQDIVPLGVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFRRPHEHF 369
DB 361 HEVQRFQDIVPLGVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFRRPHEHF 420
QY 370 LDAQHFVKPEAFIPFSAGRRACLGELPLARMELFLFTSLLOHFSFVSPTGQPRSHHG 429
DB 421 LDAQHFVKPEAFIPFSAGRRACLGELPLARMELFLFTSLLOHFSFVSPTGQPRSHHG 480
QY 430 FAFLVTPSPYELCAVPR 446
DB 481 FAFLVTPSPYELCAVPR 497

RESULT 2
G02938
probable debrisoquine 4-hydroxylase (EC 1.14.14.-) cytochrome P450 - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Mar-2000
C:Accession: G02938
R:Lawton, M.P.; Laddison, K.J.; Speirs, A.A.; Mankowski, D.C.; Tweedie, D.J.
submitted to the EMBL Data Library, October 1995
A:Reference number: G12616
A:Accession: G02938
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-497 <LAW>
A:Cross-references: EMBL:U38218; NID:g1022899; PIDN:AAA79722.1; PID:g1022900
C:Genetics:
A:Gene: CYP2D17
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; heme; iron; metalloprotein; monooxygenase; oxidoreductase; th
F:302-465/Domain: cytochrome P450 homology <P45>
F:443/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 91.9%; Score 2140.5; DB 1; Length 497;
Best Local Similarity 83.9%; Pred. No. 2.7e-157;
Matches 417; Conservative 9; Mismatches 20; Indels 51; Gaps 1;

QY 1 MGLEALVPLAVIAIFLLLVLMHRRQWAAARYSPGPLPLGGLNLLHVDVFQNTYCFDQ 60

DB 1 MELDALVPLAVIAIFLLLVLMHRRQWAAARYSPGPLPLGGLNLLHVDVFQNTYCFDQ 60
QY 61 LRRRGDVFSIQAWTPVVLNGLAAVREALVTHGEDTADRPVPIITQILGFGPSQ--- 117
DB 61 LRRRGDVFSIQAWTPVVLNGLAAVREALVTHGEDTADRPVPIITQILGFGPSQGVF 120
QY 118 -----GRPRPNGLLDK 129
DB 121 LARYGPWREORRFSVSTLRNLGLGKKSLEQWVTEAAACLAAPNHSGRPRPNGLLDK 180
QY 130 AVSNVIASITCGRRFEYDDPRFLRLDLAQGLKEESGFLREVLANVPLVLLHIPALAGKV 189
DB 181 AVSNVIASITCGRRFEYDDPRFLRLDLAQGLKEESGFLREVLANVPLVLLHIPALAGKV 240
QY 190 LRFQKAFITQDLDELTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPESFNDENRIVVA 249
DB 241 LRFQKAFITQDLDELTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPESFNDENRIVVA 300
QY 250 DLFSAGMVTSTTLAWGLLMLLHPDVQRRVQOEIDDDVIGVRRPEMGDOAHMPYTTAVI 309
DB 301 DLFSAGMVTSTTLAWGLLMLLHPDVQRRVQOEIDDDVIGVRRPEMGDOAHMPYTTAVI 360
QY 310 HEVQRFQDIVPLGVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFRRPHEHF 369
DB 361 HEVQRFQDIVPLGVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPFRRPHEHF 420
QY 370 LDAQHFVKPEAFIPFSAGRRACLGELPLARMELFLFTSLLOHFSFVSPTGQPRSHHG 429
DB 421 LDAQHFVKPEAFIPFSAGRRACLGELPLARMELFLFTSLLOHFSFVSPTGQPRSHHG 480
QY 430 FAFLVTPSPYELCAVPR 446
DB 481 FAFLVTPSPYELCAVPR 497

RESULT 3
S37284
cytochrome P450 2D - bovine
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Mar-2000
C:Accession: S37284; S29295; S29862
R:Tsunooka, Y.; Matsuo, Y.; Higuchi, R.; Ichikawa, Y.
Eur. J. Biochem. 208, 739-746, 1992
A:Title: Characterization of the cytochrome P-450IID subfamily in bovine liver. Nucleoti
A:Reference number: S29295; MUID:93011103; PMID:1396678
A:Accession: S37284
A:Molecule type: mRNA
A:Residues: 1-500 <TSU>
A:Cross-references: EMBL:X68481; NID:g295; PIDN:CAA48501.1; PID:g296
A:Experimental source: clone pBVL 180
A:Accession: S29295
A:Molecule type: mRNA
A:Residues: 14-111, 'R', 113-131, 'R', 133-162, 'L', 164-178, 'G', 180-219, 'P', 221-247, 'R', 249-2
A:Cross-references: EMBL:X68013; NID:g293; PIDN:CAA48149.1; PID:g294
A:Experimental source: clone pBV 76
C:Genetics:
A:Gene: CYP2D
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F:305-468/Domain: cytochrome P450 homology <P45>
F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 75.9%; Score 1769; DB 1; Length 500;
Best Local Similarity 69.6%; Pred. No. 1.2e-128;
Matches 348; Conservative 39; Mismatches 59; Indels 54; Gaps 2;

QY 1 MGL-----EALVPLAVIAIFLLLVLMHRRQWAAARYSPGPLPLGGLNLLHVDVFQNTYCF 57
DB 1 MGLSGDTGLAVALLIFLLLDLMHRRSRWAPRPPGPTPLPLVGLNLLQVDFDPRPS 60
QY 58 FDLRRRFGDVFSIQAWTPVVLNGLAAVREALVTHGEDTADRPVPIITQILGFGPSQ 117

```
Db 61 ENQRRRFGNVFSLOQVTPVVVVLNGLAAVREALVYRSQTDADRPFPVAYEHLGYGPRAE 120
QY 118 -----GRPPRPNGL 126
Db 121 GVILARYGDAAWQRFPSTLTTLNFGLGKKSLEQWTEBASCSAAFAQAGPFPSPMDL 180
QY 127 LKAVSNVIASTLTCGRFFEYDDPRFLRLDOLAQEGKKEBSGFLREVNLNAVPLLHHPALA 186
Db 181 LNKAVSNVIASTLTCGRFFEYDDPRFLRLDOLAQEGKKEBSGFLREVNLNAVPLLHHPALA 240
QY 187 GKVLRFQKAFATOLDDELTEHRTMTDPAOPPRDLTEAFIAEMEKAKGNPSESSFNENLRI 246
Db 241 ARVFPQAKFAMALIDELAEQKTRDTPQPRHLTDFAFDEVKEAKGNPSESSFNENLRL 300
QY 247 VVADLFSAGMVTSTTTLAWGLLMLHPDVQRRVQOEIDDDVIGQVRRPMPGQAHMPYTT 306
Db 301 VVADLFSAGMVTSTTTLAWGLLMLHPDVQRRVQOEIDDDVIGQVRRPMPGQAHMPYTT 360
QY 307 AVIHEVQRPDIPVLPVTHMTSRDIEVQGFRIIPKGTTLITNLSVLKDAVWEKPPRPH 366
Db 361 AVVHEVQRPADIIVPLGLPHTSRDIEVQGFRIIPKGTTLITNLSVLKDETVWEKPPRPH 420
QY 367 EHFLDAQGHFVKPEAFIPFSAGRACLGEPFLARMELFLFTSLLOHFSFVPTGQPRPSH 426
Db 421 EHFLDAQGHFVKPEAFIPFSAGRACLGEPFLARMELFLFTSLLOHFSFVPTGQPRPSH 480
QY 427 HGVFAFLVTPSPYELCAVPR 446
Db 481 HGVFAFLVTPAPYQLCAVPR 500

RESULT 4
JE0258
cytochrome P450 2D23 - rabbit
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Mar-2000
C:Accession: JE0258
R:Yamamoto, Y.; Ishizuka, M.; Takada, A.; Fujita, S.
J. Biochem. 124, 503-508, 1998
A:Title: Cloning, tissue distribution, and functional expression of two novel rabbit cytochrome P450s
A:Reference number: JE0258; MUID:98391821; PMID:9722658
A:Accession: JE0258
A:Molecule type: mRNA
A:Residues: 1-500 <YAM>
A:Cross-references: DDBJ:AB008784
A:Experimental source: liver
C:Comment: This protein shows high drug metabolizing activity.
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metalloprotein
F:9-25/Domain: transmembrane #status predicted <TM1>
F:305-468/Domain: cytochrome P450 homology <CYP>
F:310-326/Domain: transmembrane #status predicted <TM2>
F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 75.2%; Score 1751; DB 1; Length 500;
Best Local Similarity 69.6%; Pred. No. 3e-127;
Matches 348; Conservative 31; Mismatches 67; Indels 54; Gaps 2;

QY 1 MGL---EALVPLAVIAIFLLVDMHRRQRWAARYSPGFLPLGGLNGLLHVDFQNTPYC 57
Db 1 MGLSGLAFLAVIAIFLLVDMHRRQRWAARYSPGFLPLGGLNGLLHVDFQNTPYC 60
QY 58 FQLRRRFGDVSFLQAVTPVVVVLNGLAAVREALVTHGSDTADRPVPTIQLIGRGRSQ 117
Db 61 FQLRRRYGDVSFLQAVTPVVVVLNGLAAVREALVTHGSDTADRPVPTIQLIGRGRSQ 120
QY 118 -----GRPPRPNGL 126
Db 121 GVVMARYGPANWRQRFPSTLTNFGLGKKSLEQWTEBATCLCAAFADHAGCPSPSML 180
QY 127 LKAVSNVIASTLTCGRFFEYDDPRFLRLDOLAQEGKKEBSGFLREVNLNAVPLLHHPALA 186
```

```
Db 181 LNKAVCNVIASTLTCGRFFEYDDHRLTFLMDLTQTILKESTGNLPQVLNVIPIILLRPLGV 240
QY 187 GKVLRFQKAFATOLDDELTEHRTMTDPAOPPRDLTEAFIAEMEKAKGNPSESSFNENLRI 246
Db 241 DKVFGQKAFMALLDELTEHRTMTDPAOPPRDLTDFAFDDQVEKAKGNPSESSFNENLRL 300
QY 247 VVADLFSAGMVTSTTTLAWGLLMLHPDVQRRVQOEIDDDVIGQVRRPMPGQAHMPYTT 306
Db 301 VVTDLFAAGMVTSTTSLWALLMLHPDVQRRVQOEIDDEVIGPARRPMPGQAHMPYTT 360
QY 307 AVIHEVQRPDIPVLPVTHMTSRDIEVQGFRIIPKGTTLITNLSVLKDAVWEKPPRPH 366
Db 361 AVVHEVQRPADIIVPLGVPHTSRDIEVQGFRIIPKGTTLITNLSVLKDAVWEKPPRPH 420
QY 367 EHFLDAQGHFVKPEAFIPFSAGRACLGEPFLARMELFLFTSLLOHFSFVPTGQPRPSH 426
Db 421 EHFLDAQGHFVKPEAFIPFSAGRACLGEPFLARMELFLFTSLLOHFSFVPTGQPRPSH 480
QY 427 HGVFAFLVTPSPYELCAVPR 446
Db 481 QGAPATLVTPAPYQLCAVPR 500

RESULT 5
JC5819
cytochrome P450 2D [validated] - pig
N:Alternate names: 25-hydroxyvitamin D(3) 25-monoxygenase; cytochrome P450(14DM); cytochrome P450 2D
N:Contains: lanosterol 14 alpha-demethylase; vitamin D3 25-hydroxylase (EC 1.14.14.-)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 18-Aug-2000
C:Accession: JC5819; PC4502; S27177; S17048
R:Postlind, H.; Axen, E.; Bergman, T.; Wikvall, K.
Biochem. Biophys. Res. Commun. 241, 491-497, 1997
A:Title: Cloning, structure, and expression of a cDNA encoding vitamin D3 25-hydroxylase
A:Reference number: JC5819; MUID:98086378; PMID:9425298
A:Accession: JC5819
A:Molecule type: DNA
A:Residues: 1-500 <POS>
A:Cross-references: GB:Y16417; NID:g2956687; PIDN:CAA76205.1; PID:g2956688
A:Accession: PC4502
A:Molecule type: protein
A:Residues: 2-57;249-273;408-430 <AXE>
A:Experimental source: liver
R:Axen, E.; Bergman, T.; Wikvall, K.
Biochem. J. 287, 725-731, 1992
A:Title: Purification and characterization of a vitamin D(3) 25-hydroxylase from pig liver
A:Reference number: S27177; MUID:93075023; PMID:1445236
A:Accession: S27177
A:Molecule type: protein
A:Residues: 2-17 <AXW>
A:Experimental source: liver
R:Sono, H.; Sonoda, Y.; Sato, Y.
Biochim. Biophys. Acta 1078, 388-394, 1991
A:Title: Purification and characterization of cytochrome P-450(14DM) (lanosterol 14-alpha-hydroxylase) from pig liver
A:Reference number: S17048; MUID:91316123; PMID:1859829
A:Accession: S17048
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 2-11 <SON>
A:Note: 6-Leu was also found
C:Comment: This enzyme catalyzes the first step in the metabolic activation of vitamin D
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metalloprotein
F:305-468/Domain: cytochrome P450 homology <P45>
F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 74.6%; Score 1739; DB 1; Length 500;
Best Local Similarity 69.0%; Pred. No. 2.6e-126;
Matches 345; Conservative 40; Mismatches 61; Indels 54; Gaps 2;

QY 1 MGL---EALVPLAVIAIFLLVDMHRRQRWAARYSPGFLPLGGLNGLLHVDFQNTPYC 57
Db 1 MGLSGLAFLAVIAIFLLVDMHRRQRWAARYSPGFLPLGGLNGLLHVDFQNTPYC 60
```

Db 1 MGLLTGDLGILALAMVIFLLLVLDLMHRRSRWAPYPGMPPLPGLGNLLQVNFQDRLS 60
Qy 58 FQIARRRGDVESIQAWTPVVVINGLAARVLAALVTHGDDTADRPVPTQIILGCPRSQ 117
Db 61 FQIARRRGDVESIQAWTPVVVINGLAARVLAALVTHGDDTADRPVPTQIILGCPRSQ 120
Qy 118 -----GRPRPNGL 126
Db 121 GVILARYGKAWREQRFSVSTLRNFGLGKKSLEWVTQEAASCLCAAFADQAGRPSPNNL 180
Qy 127 LKAVSNVIASITCGRRFYDDPRFLRLDLAOGLEKESGFLREVLAAPVLLHLPALA 186
Db 181 LNKAVSNVIASITCGRRFYDDPRFLRLDLAOGLEKESGFLREVLAAPVLLHLPALA 240
Qy 187 GKVLRFQKFLTLQDELLTEHMTWDPAPPRDLTEAFIAEMKAKGNPESFNDENLRI 246
Db 241 AKFLPRQKAFVIMDELITEHMTWDPAPPRDLTEAFIAEMKAKGNPESFNDENLRL 300
Qy 247 VVADLFSAGMVTSTTLAWGLLMLHPDVQRRVQOEIDDVIGQVRRPEMGDOAHMPYTT 306
Db 301 VVADLFSAGMVTSTTLAWGLLMLHPDVQRRVQOEIDDVIGQVRRPEMGDOAHMPYTT 360
Qy 307 AVIHEVQRGDIVPLGVTHTMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPFREHP 366
Db 361 AVLHEVQRGDIVPLGVTHTMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPFREHP 420
Qy 367 EHFLDAQGHVQPEAFPLPSAGRRACLGEPFLARMELFFFTSLQHFSSVPTGQPRSSH 426
Db 421 EHFLDAQGHVQPEAFPLPSAGRRACLGEPFLARMELFFFTSLQHFSSVPTGQPRSSH 480
Qy 427 HGVFAFLVTPSPYELCAVR 446
Db 481 HGVFAFLVTPSPYELCAVR 500
RESULT 6
D31579
Cytochrome P450 2D4 - rat
N:Alternate names: cytochrome P450C2F3
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: S16873; D31579; I52313
R:Matsumura, E.; Umeno, M.; Gonzalez, F.J.
J. Mol. Evol. 30, 155-169, 1990
A:Title: The rat P450 IID subfamily: complete sequences of four closely linked genes and 50 active site.
A:Reference number: S16871; MUID:90189185; PMID:2107330
A:Accession: S16873
A:Molecule type: DNA
A:Residues: 1-500 <MAT>
A:CROSS-references: EMBL:X52029; NID:g57815; PIDN:CAA36271.1; PID:g57816
R:Ishida, N.; Tawaragi, Y.; Inuzuka, C.; Sugita, O.; Kubota, I.; Nakazato, H.; Noguchi, R. Biochem. Biophys. Res. Commun. 156, 681-688, 1988
A:Title: Four species of cDNAs for cytochrome P450 isozymes immunorelated to P450C-M/P e
A:Reference number: A90151; MUID:89050091; PMID:3190674
A:Accession: D31579
A:Molecule type: mRNA
A:Residues: 177-500 <ISH>
A:CROSS-references: EMBL:M22331; NID:g203829; PIDN:AAA41052.1; PID:g203830
R:Kawashima, H.; Strobel, H.W.
Biochem. Biophys. Res. Commun. 209, 535-540, 1995
A:Title: cDNA cloning of a novel rat brain cytochrome P450 belonging to the CYP2D subfam
A:Reference number: I52313; MUID:95251650; PMID:7733922
A:Accession: I52313
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-326, 'R', 328-399, 'I', 401-472, 'A', 474-479, 'N', 481-482, 'V', 484-500 <RES>
A:CROSS-references: GB:S77859; NID:g1200515; PIDN:AC52882.1; PID:g1200516
A:Experimental source: brain, strain Sprague-Dawley
C:Genetics:
A:Gene: CYP2D4
A:Introns: 63/3; 121/1; 172/1; 225/3; 284/3; 332/1; 394/3; 442/1

C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F:9-25/Domain: transmembrane #status predicted <TM1>
F:305-468/Domain: cytochrome P450 homology <CYP>
F:310-326/Domain: transmembrane #status predicted <TM2>
F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted
Query Match 74.0%; Score 1724.5; DB 1; Length 500;
Best Local Similarity 68.9%; Pred. No. 3.3e-125;
Matches 339; Conservative 33; Mismatches 69; Indels 51; Gaps 1;
Qy 6 LVPALAVIAIFLLVLDLMHRRQRWAARYSPGLPLPGLGNLLHVDVFQNTPYCFDQRRRF 65
Db 9 LWPALAVIAIFLLVLDLMHRRQRWTSRYPPGVPWVGLNLLQIDFQNPAGFQKLRCP 68
Qy 66 GDFSLQAWTPVVVINGLAARVLAALVTHGDDTADRPVPTQIILGCPRSQ----- 117
Db 69 GDFSLQAFESVVVINGLPAALREALVKYSDDTADRPPLHFNDQSGFGPRSGVVLAARYG 128
Qy 118 -----GRPRPNGLDKAVSNV 134
Db 129 PAWRQRRFSVSTFRHFGLGKKSLEWVTBACCLCAAFADHSGFPFSPNTLLDKAVCNV 188
Qy 135 IASLTGCGRRFYDDPRFLRLDLAOGLEKESGFLREVLAAPVLLHLPALAGKVLRFQK 194
Db 189 IASLTGCGRRFYDDPRFLRLDLAOGLEKESGFLREVLAAPVLLHLPALAGKVLRFQK 248
Qy 195 AFLTQDELLTEHMTWDPAPPRDLTEAFIAEMKAKGNPESFNDENLRIVVADLPSA 254
Db 249 AFVAMLDLLEHMTWDPAPPRDLTEAFIAEMKAKGNPESFNDENLRIVVADLPSA 308
Qy 255 GMTVSTTLANGLLMLHPDVQRRVQOEIDDVIGQVRRPEMGDOAHMPYTTAVIHEVQR 314
Db 309 GMTVSTTLANGLLMLHPDVQRRVQOEIDDVIGQVRRPEMGDOAHMPYTTAVIHEVQR 368
Qy 315 FGDIVPLGVTHTMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPFREHPFELDAQG 374
Db 369 FADILPLGVPKHTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPFREHPFELDAQG 428
Qy 375 HFVQPEAFPLPSAGRRACLGEPFLARMELFFFTSLQHFSSVPTGQPRSSHGVFAFLV 434
Db 429 NFVQHEAFMPFSAGRRACLGEPFLARMELFFFTSLQHFSSVPTGQPRSSHGVFAFLV 488
Qy 435 TPSPYELCAVR 446
Db 489 TPSPYELCAVR 500
RESULT 7
J50259
Cytochrome P450 2D24 - rabbit
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Mar-2000
C:Accession: J50259
R:Yamamoto, Y.; Ishizuka, M.; Takada, A.; Fujita, S.
J. Biochem. 124, 503-508, 1998
A:Title: Cloning, tissue distribution, and functional expression of two novel rabbit cyt
A:Reference number: J50258; MUID:98391821; PMID:9722658
A:Accession: J50259
A:Molecule type: mRNA
A:Residues: 1-500 <YAM>
A:CROSS-references: DBJ:AB008785
C:Comment: This protein shows high drug metabolizing activity.
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F:9-25/Domain: transmembrane #status predicted <TM1>
F:305-468/Domain: cytochrome P450 homology <CYP>
F:310-326/Domain: transmembrane #status predicted <TM2>
F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted
Query Match 72.8%; Score 1697; DB 1; Length 500;
Best Local Similarity 68.0%; Pred. No. 4.4e-123;

C;Accession: S09611; A32970; S16874; B31579
R;Ishida, N.; Inuzuka, C.; Tawaragi, Y.; Sugita, O.; Nakazato, H.; Noguchi, T.; Sassa, S.
Nucleic Acids Res. 17, 6407, 1989
A;Title: Cytochrome P450CMF cDNA: nucleotide sequence of P450CMF1b.
A;Reference number: S09611; MUID:89366685; PMID:2771656
A;Accession: S09611
A;Molecule type: mRNA
A;Residues: 1-504 <ISH>
A;Cross-references: EMBL:M25143; NID:G203775; PIDN:AAA41034.1; PID:G203775
R;Matsunaga, E.; Zanger, U.M.; Hardwick, J.P.; Gelboin, H.V.; Meyer, U.A.; Gonzalez, F.J.
Biochemistry 28, 7349-7355, 1989
A;Title: The CYP2D gene subfamily: analysis of the molecular basis of the debrisoquine 4-
A;Reference number: A32970; MUID:90057430; PMID:2819073
A;Accession: A32970
A;Molecule type: mRNA
A;Residues: 1-504 <MAT1>
A;Cross-references: EMBL:J02869; NID:G203673; PIDN:AAA41003.1; PID:G203674
R;Matsunaga, E.; Umeno, M.; Gonzalez, F.J.
J. Mol. Evol. 30, 155-169, 1990
A;Title: The rat P450 IID subfamily: complete sequences of four closely linked genes and
50 active site.
A;Reference number: S16871; MUID:90189185; PMID:2107330
A;Accession: S16874
A;Molecule type: DNA
A;Residues: 1-504 <MAT2>
A;Cross-references: EMBL:X52030; NID:G57817; PIDN:CAA36272.1; PID:G57818
R;Ishida, N.; Tawaragi, Y.; Inuzuka, C.; Sugita, O.; Kubota, I.; Nakazato, H.; Noguchi, F.
Biochem. Biophys. Res. Commun. 156, 681-688, 1988
A;Title: Four species of cDNAs for cytochrome P450 isozymes immunorelated to P450C-M/F
A;Reference number: A90151; MUID:89050091; PMID:3190674
A;Accession: B31579
A;Molecule type: mRNA
A;Residues: 18-504 <IS2>
A;Cross-references: EMBL:M22329; NID:G203806; PIDN:AAA41045.1; PID:G203807
C;Genetics:
A;Gene: CYP2D5
A;Introns: 63/3; 121/1; 172/1; 225/3; 284/3; 332/1; 394/3; 442/1
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F;305-468/Domain: cytochrome P450 homology <P45>
F;446/Binding site: heme iron (Cys) (axial ligand) #status predicted
Query Match 69.3%; Score 1614.5; DB 1; Length 504;
Best Local Similarity 64.6%; Pred. No. 1e-116;
Matches 318; Conservative 38; Mismatches 85; Indels 51; Gaps 1;
QY 6 LVPLAVIVAIFLLVLDLHRRQRWAARYSGPLPLPGLGNLLHVDFTQNTPCFDQLRRRF 65
DB 9 LWPMAIFTVIFILLVLDLHRRQWTSKYPGPVWPVLGNLLQVDPSPNMPYSMYKLQHRF 68
QY 66 GDVFSIQLAWTPVVVLNGLAAVREALVTHGEDTADRPVPVPIQILGFGPRSQ----- 117
DB 69 GDVFSIQMGKPMVIVNRLKAVQEVLTGHEDTADRPVPVPIKCLGVKPRSQGVFPASVG 128
QY 118 -----GRPFPRNGLLDKAVSNV 134
DB 129 PEWREQRRESVSLRTFTGMRKSLSEBWTKEAGHLCDAFTAQNGRSINPKAMLNKALCNV 188
QY 135 IASLTGCRPEYDDPRLRLDLAOLKEESFLAEVLNVPVLLHHPALAGKVLRFQK 194
DB 189 IASLIFARPEYEDPVLRLMLTUVESLSEVSGFIEVLNTPFALLRIFGLADKVFQGQK 248
QY 195 AFLTQDLELLTHRTWTDPAQPPDLTEAFIAMEKAKGNPSSFNDELIRIVADLFS 254
DB 249 TFNAFLDNLAAENRTTWDPAQPPNLTDAFLAEVAKGNPSSFNDELIRVWVDLFTA 308
QY 255 GMYTTSTLAWGLLMLHPDVQRVQOEIDVIGVRRPEMGDQAHMPTTAVIHEVQR 314
DB 309 GMYTTATTLTALLMLLYPDVQRVQOEIDEVIGVRRPEMTDQAHMPTTAVIHEVQR 368
QY 315 FGDIVPLGVTHMSTRDIEVQGFPIPKGTTLITNLSVLKDEAVWEKPFPHFHLDAQG 374
DB 369 FGDIAFLNLPRTSCDIEVQDFVPIKGTTLITNLSVLKDETVWEKPLRPHFHLDAQG 428

QY 375 HFVKPEAFIPFSGARRACIGELPLARMELFTTLLQHFSSVPTGQPRSHGVFAFLV 434
DB 429 NFVKHEAFMPFSGARRACIGELPLARMELFTTLLQHFSSVPTGQPRSPSTILGNAISV 488
QY 435 TPSPYELCAVPR 446
DB 489 APLPYOLCAVR 500

RESULT 12

A26822
debrisoquine 4-hydroxylase (EC 1.14.14.14.-) cytochrome P450 2D1 - rat
N;Alternate names: cytochrome P450 UT-7; cytochrome P450db1
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 03-Mar-2000
C;Accession: A26822; A30495; B32970; C32970; A31579; JCA158; S39761
R;Gonzalez, F.J.; Matsunaga, T.; Nagata, K.; Meyer, U.A.; Nebert, D.W.; Pastewka, J.; K
DNA 6, 149-161, 1987
A;Title: Debrisoquine 4-hydroxylase: characterization of a new P450 gene subfamily, reg
A;Reference number: A90957; MUID:87217961; PMID:3582092
A;Accession: A26822
A;Molecule type: mRNA
A;Residues: 1-504 <CON>
A;Cross-references: EMBL:M16654; NID:G203833; PIDN:AAA41054.1; PID:G203834
A;Accession: A30495
A;Molecule type: protein
A;Residues: 'X', '5', '7', 'X', '9', 'XX', '12-23 <GO2>
R;Matsunaga, E.; Zanger, U.M.; Hardwick, J.P.; Gelboin, H.V.; Meyer, U.A.; Gonzalez, F.
Biochemistry 28, 7349-7355, 1989
A;Title: The CYP2D gene subfamily: analysis of the molecular basis of the debrisoquine
A;Reference number: A32970; MUID:90057430; PMID:2819073
A;Accession: B32970
A;Molecule type: mRNA
A;Residues: 1-504 <MAL>
A;Cross-references: EMBL:J02867; NID:G203669; PIDN:AAA41001.1; PID:G203670
A;Accession: C32970
A;Molecule type: mRNA
A;Residues: 1-122, 'VF', '125-172, 'R', '174-379, 'I', '381-504 <MAT1>
R;Ishida, N.; Tawaragi, Y.; Inuzuka, C.; Sugita, O.; Kubota, I.; Nakazato, H.; Noguchi, F.
Biochem. Biophys. Res. Commun. 156, 681-688, 1988
A;Title: Four species of cDNAs for cytochrome P450 isozymes immunorelated to P450C-M/F
A;Reference number: A90151; MUID:89050091; PMID:3190674
A;Accession: A31579
A;Molecule type: mRNA
A;Residues: 1-122, 'VF', '125-172, 'R', '174-379, 'I', '381-504 <ISH>
A;Cross-references: EMBL:M22328; NID:G203802; PIDN:AAA41043.1; PID:G203803
R;Jiang, Q.; Voigt, J.M.; Colby, H.D.
Biochem. Biophys. Res. Commun. 209, 1149-1156, 1995
A;Title: Molecular cloning and sequencing of a guinea pig cytochrome P4502D (CYP2D16) :
A;Reference number: JCA153; MUID:95251703; PMID:7733969
A;Accession: JCA158
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-504 <JIA>
R;Ohishi, N.; Imaoka, S.; Suzuki, T.; Funae, Y.
Biochim. Biophys. Acta 1158, 227-236, 1993
A;Title: Characterization of two P-450 isozymes placed in the rat CYP2D subfamily.
A;Reference number: S39761; MUID:94072607; PMID:8251521
A;Accession: S39761
A;Molecule type: protein
A;Residues: 1-9, 'X', '11-13 <OHI>
C;Genetics:
A;Gene: CYP2D1
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metall
F;1-504/Product: cytochrome P450 2D1 #status experimental <MAT>
F;4-504/Product: cytochrome P450 2Div #status experimental <MAT2>
F;9-25/Domain: transmembrane #status predicted <TM1>
F;305-468/Domain: cytochrome P450 homology <CYP>
F;310-326/Domain: transmembrane #status predicted <TM2>
F;446/Binding site: heme iron (Cys) (axial ligand) #status predicted

F:305-468/Domain: cytochrome P450 homology <CYP>
F:310-326/Domain: transmembrane #status predicted <TM2>
F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 68.6%; Score 1597.5; DB 1; Length 504;
Best Local Similarity 62.4%; Pred. No. 2.1e-115;
Matches 307; Conservative 54; Mismatches 80; Indels 51; Gaps 1;

QY 6 LVPLAVIAIFLLVLDLHRRORWAARYSPGLPLGLGNLLHVDQNTFYCFDQLRRRP 65
DB 9 LWPVAIFTWIFILLVLDLTHQQRWTSRYPPGPWPVGLGNLLQVLDGNMPSLYLKQRY 68

QY 66 GDVFSIQLAWTPVVVINGLAAREALVTHGEDTADRPVPVITQILGFGRRSQ----- 117
DB 69 GDVFSIQMAWKPIVINGLAKKEMLLTCGEDTADRPVPVIFEYLGKPGSQGVWLAPYG 128

QY 118 -----GRPFRNGLLDKAVSNV 134
DB 129 PEWRQRRFSVTLNRFGLGKKSLEDVWTKAEANHLCDATAQAAGQPINPNMLNKSTCNV 188

QY 135 IASLTGRRFEYDDPRFLRLDLAQEGKKEESGFLREVNLNAVPLVLLHIIPALAGKVLRFQK 194
DB 189 IASLIAPARRFEYEDPFLIRMLKVLQESLTSGLIPEVLNAPFILLIRIPLADKALQGGK 248

QY 195 AFLTQDELLETHRTWTDPAQPRDLTEAFIAEMEKAKGNPSSFNDENLRIIVADLFSA 254
DB 249 SFNAILDNLNLTENRTTWDVQAPRNLTDAFLAEIEKAKGNPSSFNDENLRIIVRDLFGA 308

QY 255 GMVTTSTTLAWGLLMLHPDVQRRVQOEIDDVIGOVRRPEMGDQAHMPYTTAVIHEVQR 314
DB 309 GMVTTSTTLAWGLLMLHPDVQRRVQOEIDDVIGOVRRPEMGDQAHMPYTTAVIHEVQR 368

QY 315 FGDIVPLGVTHMTSRDIEVQGFRIIPKGTTLITNLSSVLKDEAVWEKPFHFHFLDAQG 374
DB 369 FGDIVPVNLPRTSHDIEVQDFLIIPKGTILLPNMSSMLKDETVWEKPLFHFHFLDAQG 428

QY 375 HFVKPEAFLPFSAGRRACGEPLARMEFLFTTSLLOHFSFVSPTGQPRPSHHGVFAFLV 434
DB 429 HFVKPEAFMPFSAGRRACGEPLARMEFLFTTCLLQRFVSFVDPGQPPSPNSGVYGLV 488

QY 435 TSPSYELCAVPR 446
DB 489 APSYQLCAVIR 500

RESULT 15
I49427
Cytochrome P450 16a-mb1 - western wild mouse
C:Species: Mus spretus (western wild mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 04-Mar-2000
C:Accession: I49427
R:Sueyoshi, T.; Kobayashi, R.; Nishio, K.; Aida, K.; Moore, R.; Wada, T.; Handa, H.; Neg
Mol. Cell. Biol. 15, 4158-4166, 1995
A:Title: A nuclear factor (NF2d9) that binds to the male-specific P450 (Cyp 2d-9) gene i
A:Reference number: A57454; MUID:95349581; PMID:7623810
A:Accession: I49427
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-504.<RES>
A:Cross-references: EMBL:U20087; NID:g951099; PID:AAC52245.1; PID:g951100
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; heme; iron; metalloprotein; transmembrane protein
F:305-468/Domain: cytochrome P450 homology <P45>
F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 67.8%; Score 1580.5; DB 2; Length 504;
Best Local Similarity 62.8%; Pred. No. 4.3e-114;
Matches 309; Conservative 48; Mismatches 84; Indels 51; Gaps 1;

QY 6 LVPLAVIAIFLLVLDLHRRORWAARYSPGLPLGLGNLLHVDQNTFYCFDQLRRRP 65
DB 9 LWPVAIFTWIFILLVLDLTHQQRWTSRYPPGPWPVGLGNLLQVLDGNMPSLYLKQRY 68

Search completed: February 25, 2004, 04:31:31
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 03:09:36 ; Search time 18 Seconds
(without alignments)
1290.183 Million cell updates/sec

Title: US-09-820-788A-2

Perfect score: 2330

Sequence: 1 MGLEALVPLAIVAFLLLV.....HGVAFLVTPSPYELCAVPR 446

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2282.5	98.0	497	1 CPD6 HUMAN	P10635 homo sapien
2	2140.5	91.9	497	1 CPDH MACFA	Q29488 macaca fasc
3	2069.5	88.8	497	1 CPDU CALJA	O18992 callithrix
4	1736.5	74.5	499	1 CPDP PIG	O46658 sus scrofa
5	1734.5	74.4	487	1 CPDE BOVIN	Q01361 bos taurus
6	1724.5	74.0	500	1 CPDA RAT	P13108 rattus norv
7	1708.5	73.3	500	1 CPDI RAT	O64680 rattus norv
8	1686.5	72.4	499	1 CPDE CANFA	Q29473 canis faml
9	1653.5	71.0	500	1 CPDK MESAU	O9QY95 mesocricetu
10	1642.5	70.5	500	1 CPDR MESAU	O9QY96 mesocricetu
11	1640	70.4	500	1 CPD2 RAT	P10634 rattus norv
12	1638.5	70.3	500	1 CPD3 RAT	P12938 rattus norv
13	1614.5	69.3	504	1 CPD1 RAT	P10633 rattus norv
14	1614.5	69.3	504	1 CPD5 RAT	P12939 rattus norv
15	1611	69.1	500	1 CPDG CAVPO	O64403 cavia porce
16	1597.5	68.6	504	1 CPD9 MOUSE	P11714 mus musculu
17	1578.5	67.7	504	1 CPDA MOUSE	P24456 mus musculu
18	1545.5	66.3	500	1 CPDS MESAU	O9QUJ1 mesocricetu
19	1456	62.5	505	1 CPDB MOUSE	P24457 mus musculu
20	841	36.1	501	1 CPD6 MOUSE	O54750 mus musculu
21	833.5	35.8	502	1 CPD2 HUMAN	P51589 homo sapien
22	828.5	35.6	502	1 CPJ3 RAT	P51590 rattus norv
23	813	34.9	501	1 CPJ5 MOUSE	O54749 mus musculu
24	810.5	34.8	491	1 CPB4 RABIT	P00178 oryctolagus
25	810.5	34.8	491	1 CPB5 RABIT	P12789 oryctolagus
26	808.5	34.7	491	1 CPB1 RAT	P00176 rattus norv
27	804.5	34.5	500	1 CPJ1 RABIT	P24786 oryctolagus
28	801	34.4	494	1 CPD3 RAT	P24470 rattus norv
29	800	34.3	491	1 CPB9 MOUSE	P12790 mus musculu
30	799.5	34.3	491	1 CPB2 RAT	P04167 rattus norv
31	797.5	34.2	490	1 CPD8 HUMAN	P10632 homo sapien
32	788.5	33.8	490	1 CPD4 MOUSE	P56656 mus musculu
33	787.5	33.8	490	1 CPD7 MESAU	Q08078 mesocricetu

34	785.5	33.7	490	1 CPCK MACFA	P33262 macaca fasc
35	785	33.7	490	1 CPDQ MESAU	P33263 mesocricetu
36	783	33.6	490	1 CPCT MOUSE	O64458 mus musculu
37	780	33.5	500	1 CPGB RAT	P08683 rattus norv
38	779.5	33.5	493	1 CPB1 HUMAN	P05181 homo sapien
39	779.5	33.5	504	1 CPK1 ONCMY	Q92090 oncorhynch
40	779	33.4	491	1 C2F2 MOUSE	P33267 mus musculu
41	778	33.4	487	1 CPCL CANFA	P56594 canis faml
42	778	33.4	492	1 CPBC RAT	P33272 rattus norv
43	776.5	33.3	491	1 CPB6 HUMAN	P20813 homo sapien
44	775	33.3	490	1 CPD8 MESAU	P33264 mesocricetu
45	775	33.3	492	1 CPB3 MOUSE	O55071 mus musculu

ALIGNMENTS

RESULT 1
CPD6_HUMAN
ID CPD6_HUMAN STANDARD; PRT; 497 AA.
AC P10635; Q16752;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome P450 2D6 (EC 1.14.14.1) (CYP1D6) (P450-DB1) (Debrisoquine 4-hydroxylase).
GN CYP2D6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88314109; PubMed=3410476;
RA Gonzalez F.J., Vilbois F., Hardwick J.P., McBride O.W.,
RA Nebert D.W., Gelboin H.V., Meyer U.A.;
RT "Human debrisoquine 4-hydroxylase (P4501D1): cDNA and deduced amino acid sequence and assignment of the CYP2D locus to chromosome 22.";
RL Genomics 2:174-179(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=88122614; PubMed=3123997;
RA Gonzalez F.J., Skoda R.C., Kimura S., Umeno M., Zanger U.M.,
RA Nebert D.W., Gelboin H.V., Hardwick J.P., Meyer U.A.;
RT "Characterization of the common genetic defect in humans deficient in debrisoquine metabolism.";
RL Nature 331:442-446(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90072069; PubMed=2574001;
RA Kimura S., Umeno M., Skoda R.C., Meyer U.A., Gonzalez F.J.;
RT "The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene.";
RL Am. J. Hum. Genet. 45:889-904(1989).
RN [4]
RP VARIANT LYS-281 DEL (CYP2D6*9).
RX MEDLINE=93244880; PubMed=1844820;
RA Tyndale R., Aoyama T., Broly F., Matsunaga T., Inaba T., Kalow W., Gelboin H.V., Meyer U.A., Gonzalez F.J.;
RT "Identification of a new variant CYP2D6 allele lacking the codon encoding Lys-281: possible association with the poor metabolizer phenotype.";
RL Pharmacogenetics 1:26-32(1991).
RN [5]
RP VARIANTS SER-34 AND THR-486 (CYP2D6*10).
RX MEDLINE=94115362; PubMed=8287064;
RA Yokota H., Tamura S., Furuya H., Kimura S., Watanabe M., Kanazawa I., Kondo I., Gonzalez F.J.;
RT "Evidence for a new variant CYP2D6 allele CYP2D6J in a Japanese population associated with lower in vivo rates of sparteine metabolism.";

Pharmacogenetics 3:256-263(1993).
[6] VARIANT PRO-324 (CYP2D6*7).
RX MEDLINE=95147995; PubMed=7845481;
RA Evert B., Griese E.U., Bichelbaum M.;
RT "A missense mutation in exon 6 of the CYP2D6 gene leading to a
RT histidine 324 to proline exchange is associated with the poor
RT metabolizer phenotype of sparteine.";
RL Naunyn Schmiedeberg Arch. Pharmacol. 350:434-439(1994).
[7] VARIANT GLU-212 (CYP2D6*6B/6C).
RP MEDLINE=95172594; PubMed=7868129;
RX Dally A.K., Leathart J.B., London S.J., Idle J.R.;
RA "An inactive cytochrome P450 CYP2D6 allele containing a deletion and a
RT base substitution.";
RL Hum. Genet. 95:337-341(1995).
[8] VARIANT ILE-107 (CYP2D6*17).
RP MEDLINE=97126511; PubMed=8971426;
RX Masimirembwa C., Persson I., Bertilsson L., Hasler J.,
RA Ingelman-Sundberg M.;
RT "A novel mutant variant of the CYP2D6 gene (CYP2D6*17) common in a
RT black African population: association with diminished debrisoquine
RT hydroxylase activity.";
RL Br. J. Clin. Pharmacol. 42:713-719(1996).
[9] VARIANT ARG-42 (CYP2D6*12).
RP MEDLINE=96209916; PubMed=8655150;
RX Marez D., Legrand M., Sabbagh N., Lo-Guidice J.M., Boone P., Broly F.;
RA "An additional allelic variant of the CYP2D6 gene causing impaired
RT metabolism of sparteine.";
RL Hum. Genet. 97:668-670(1996).
[10] VARIANT ARG-169 (CYP2D6*14).
RP MEDLINE=99164054; PubMed=10064570;
RX Wang S.L., Lai M.D., Huang J.D.;
RA "G169R mutation diminishes the metabolic activity of CYP2D6 in
RT Chinese.";
RL Drug Metab. Dispos. 27:385-388(1999).
CC -!- FUNCTION: Responsible for the metabolism of many drugs and
CC environmental chemicals that it oxidizes. It is involved in the
CC metabolism of drugs such as antiarrhythmics, adrenoceptor
CC antagonists, and tricyclic antidepressants.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: By pregnancy.
CC -!- POLYMORPHISM: Highly polymorphic. Oxidative drug metabolism by
CC CYP2D6 is characterized by two phenotypes, the extensive
CC metabolizer (EM) and poor metabolizer (PM). Of the Caucasian
CC populations of Europe and North America, 5%-10% are of the PM
CC phenotype and are unable to metabolize the antihypertensive drug
CC debrisoquine and numerous other drugs.
CC -!- POLYMORPHISM: Allele CYP2D6*7 was also known as CYP2D6E, allele
CC CYP2D6*9 as CYP2D6C, allele CYP2D6*10 as CYP2D6J, allele CYP2D6*17
CC as CYP2D6Z.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -!- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;
CC NOTE=CYP2D6 alleles;
CC WWW="http://www.imm.ki.se/cypalleles/cyp2d6.htm".

This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; M20403; AAA52153.1; -;
DR EMBL; X08006; CAA30807.1; -;
DR EMBL; M33388; AAA53500.1; -;
DR PIR; S01199; O4HUD1.
DR HSSP; P00179; 1DT6.
DR GENE; HGNC:2625; CYP2D6.
DR MIM; 124030; -;
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; P450_1.
DR PRINTS; PR01686; EP450CYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Polymorphism.
FT METAL 443 443 IRON (HEME AXIAL LIGAND)
FT VARIANT 11 11 V -> M (in allele CYP2D6*35).
FT VARIANT 26 26 R -> H (in allele CYP2D6*21).
FT VARIANT 28 28 R -> C (in allele CYP2D6*22).
FT VARIANT 34 34 P -> S (in allele CYP2D6*10 and allele
FT CYP2D6*14; poor debrisoquine metabolism).
FT VARIANT 42 42 G -> R (in allele CYP2D6*12; impaired
FT metabolism of sparteine).
FT FT FTId=VAR_001256.
FT VARIANT 85 85 A -> V (in allele CYP2D6*23).
FT VARIANT 107 107 T -> I (in allele CYP2D6*17; poor
FT debrisoquine metabolism).
FT VARIANT 169 169 G -> R (in allele CYP2D6*14; poor
FT debrisoquine metabolism).
FT VARIANT 212 212 G -> E (in allele CYP2D6*6B and allele
FT CYP2D6*6C).
FT VARIANT 237 237 A -> S (in allele CYP2D6*33).
FT VARIANT 281 281 Missing (in allele CYP2D6*9).
FT VARIANT 296 296 R -> C (in allele CYP2D6*2, allele
FT CYP2D6*12, allele CYP2D6*14 and allele
FT CYP2D6*17; dbSNP:16947).
FT VARIANT 297 297 FTId=VAR_008340.
FT VARIANT 311 311 I -> L (in allele CYP2D6*24).
FT VARIANT 311 311 S -> L (in dbSNP:1800754).
FT VARIANT 324 324 H -> P (in allele CYP2D6*7; loss of
FT activity).
FT VARIANT 343 343 FTId=VAR_008348.
FT VARIANT 369 369 R -> G (in allele CYP2D6*25).
FT VARIANT 410 410 FTId=VAR_008372.
FT VARIANT 410 410 I -> T (in allele CYP2D6*26).
FT VARIANT 410 410 FTId=VAR_008373.
FT VARIANT 486 486 E -> K (in allele CYP2D6*27).
FT VARIANT 486 486 S -> T (in allele CYP2D6*2, allele
FT CYP2D6*10, allele CYP2D6*12, allele
FT CYP2D6*14 and allele CYP2D6*17; impaired
FT metabolism of sparteine).
FT FT FTId=VAR_008341.
FT CONFLICT 374 374 M -> V (IN REF. 3).
FT SEQUENCE 497 AA; 55801 MW; 543F4D5F0DE8CDAC CRC64;

Query Match 98.0%; Score 2282.5; DB 1; Length 497;
Best Local Similarity 89.1%; Pred. No. 4e-154; 1; Indels 51; Gaps 1;
Matches 443; Conservative 2; Mismatches 1;

QY 1 MGLEALVPLAVIAIVAILLLVLMHRRQWAARYSPGFLPLPGLGNLLHVDFTQNTPYCFDQ 60
DB 1 MGLEALVPLAVIAIVAILLLVLMHRRQWAARYSPGFLPLPGLGNLLHVDFTQNTPYCFDQ 60

QY 61 LRRFGDVPSLOLAWTPVVVINGLAAREALVTHGEDTADRPVPITQILGFGPRSQ--- 117
DB 61 LRRFGDVPSLOLAWTPVVVINGLAAREALVTHGEDTADRPVPITQILGFGPRSQGVF 120

QY 118 -----GRPRPNGLLDK 129
DB 121 LARYGPANREQRFRFVSTLRNLGLGKKSLEQWTEBAACLAFAFANHSRGRFRPNGLLDK 180

QY 130 AVSNVIASITCGRRFEYDDPRFLRLDLAQEGLEKESGFLREVNLAVPVLLHHPALAGKV 189
DB 181 AVSNVIASITCGRRFEYDDPRFLRLDLAQEGLEKESGFLREVNLAVPVLLHHPALAGKV 240

QY 190 LRFQKAFLTQLDELLETHRTWMDPAQPPRDLTEAFLAEMEKAAGNPFSSFNENLRIVA 249
DB 241 LRFQKAFLTQLDELLETHRTWMDPAQPPRDLTEAFLAEMEKAAGNPFSSFNENLRIVA 300

QY 250 DLFSAGMVTSTTTLAWGLLMLHPDVQRRVQOEIDDDVIGQVRRPDMGQAHMPYTTAVI 309
DB 301 DLFSAGMVTSTTTLAWGLLMLHPDVQRRVQOEIDDDVIGQVRRPDMGQAHMPYTTAVI 360

QY 310 HEVQFGDIVPLGVTHMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPRFHPHF 369
DB 361 HEVQFGDIVPLGVTHMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPRFHPHF 420

QY 370 LDAQHFVKEAPEALPFSAGRRACLGELPLARMELFFFTSLLQHFSSVPTGQPRPSHHGV 429
DB 421 LDAQHFVKEAPEALPFSAGRRACLGELPLARMELFFFTSLLQHFSSVPTGQPRPSHHGV 480

QY 430 FAFVLTSPSYELCAVPR 446
DB 481 FAFVLTSPSYELCAVPR 497

RESULT 2

CPDH MACFA STANDARD; PRT; 497 AA.

ID CPDJ CALJA
AC Q29488;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D17 (EC 1.14.14.1) (CYP1D17).
GN CYP2D17.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Lawton M.P., Laddison K.J., Speirs A.A., Mankowski D.C., Tweedie D.J.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -! CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -! SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By
CC similarity).
CC -! SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U38218; AAA79722.1; -
DR PIR; G02938; G02938.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00667; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum
FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 497 AA; 56010 MW; 3594AA88F04E58B1 CRC64;

Query Match 91.9%; Score 2140.5; DB 1; Length 497;
Best Local Similarity 83.9%; Pred. No. 4.4e-144;
Matches 417; Conservative 9; Mismatches 20; Indels 51; Gaps 1;

QY 1 MGLEALVPLAVIAIVAILLLVLMHRRQWAARYSPGFLPLPGLGNLLHVDFTQNTPYCFDQ 60
DB 1 MELDALVPLAVIAIVAILLLVLMHRRQWAARYSPGFLPLPGLGNLLHVDFTQNTPYCFDQ 60

QY 61 LRRFGDVPSLOLAWTPVVVINGLAAREALVTHGEDTADRPVPITQILGFGPRSQ--- 117
DB 61 LRRFGDVPSLOLAWTPVVVINGLAAREALVTHGEDTADRPVPITQILGFGPRSQGVF 120

QY 118 -----GRPRPNGLLDK 129
DB 121 LARYGPANREQRFRFVSTLRNLGLGKKSLEQWTEBAACLAFAFANHSRGRFRPNGLLDK 180

QY 130 AVSNVIASITCGRRFEYDDPRFLRLDLAQEGLEKESGFLREVNLAVPVLLHHPALAGKV 189
DB 181 AVSNVIASITCGRRFEYDDPRFLRLDLAQEGLEKESGFLREVNLAVPVLLHHPALAGKV 240

QY 190 LRFQKAFLTQLDELLETHRTWMDPAQPPRDLTEAFLAEMEKAAGNPFSSFNENLRIVA 249
DB 241 LRSQKAFLTQLDELLETHRTWMDPAQPPRDLTEAFLAEMEKAAGNPFSSFNENLRIVA 300

QY 250 DLFSAGMVTSTTTLAWGLLMLHPDVQRRVQOEIDDDVIGQVRRPDMGQAHMPYTTAVI 309
DB 301 DLFSAGMVTSTTTLAWGLLMLHPDVQRRVQOEIDDDVIGQVRRPDMGQAHMPYTTAVI 360

QY 310 HEVQFGDIVPLGVTHMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPRFHPHF 369
DB 361 HEVQFGDIVPLGVTHMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPRFHPHF 420

QY 370 LDAQHFVKEAPEALPFSAGRRACLGELPLARMELFFFTSLLQHFSSVPTGQPRPSHHGV 429
DB 421 LDAQHFVKEAPEALPFSAGRRACLGELPLARMELFFFTSLLQHFSSVPTGQPRPSHHGV 480

QY 430 FAFVLTSPSYELCAVPR 446
DB 481 FAFVLTSPSYELCAVPR 497

RESULT 3
CPDJ CALJA STANDARD; PRT; 497 AA.

ID CPDJ CALJA
AC O18992;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D19 (EC 1.14.14.1) (CYP1D19) (P450 CM2D-1).
GN CYP2D19.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callitrich.
OX NCBI_TaxID=9483;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;

```

RX MEDLINE=97223367; PubMed=9056237;
RA Igarashi T., Sakuma T., Isogai M., Nagata R., Kamataki T.;
RT "Marmoset liver cytochrome P450s: study for expression and molecular
RT cloning of their cDNAs.";
CC ENVIRONMENTAL CHEMICALS THAT IT OXIDIZES.
CC -!- FUNCTION: RESPONSIBLE FOR THE METABOLISM OF MANY DRUGS AND
CC ENVIRONMENTAL CHEMICALS THAT IT OXIDIZES.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D29822; BAA22155.1; -.
CC HSP; P00179; 1DT6.
CC InterPro; IPR001128; Cytochrome P450.
CC InterPro; IPR008069; EP450_CYP2D.
CC Pfam; PF00067; p450; 1.
CC PRINTS; PR01686; EP450CYP2D.
CC PROSITE; PS00086; CYTOCHROME P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 497 AA; 55911 MW; A482ABE71B4D6CAF CRC64;
Query Match 88.8%; Score 2069.5; DB 1; Length 497;
Best Local Similarity 81.5%; Pred. No. 4.6e-139;
Matches 405; Conservative 11; Mismatches 30; Indels 51; Gaps 1;
QY 1 MGLEALVPLAVTVAIFLLVLDLHRRQRWAARYSPGLPLGGLNLLHVDVFONTPCYDQ 60
DB 1 MGLEALVPLAVTVAIFLLVLDLHRRQRWAARYSPGLPLGGLNLLHVDVFONTNSNQ 60
QY 61 LRRRGDFVSLQAWTPVVLNGLAAVREALVTHGEDTADRPVPIQILGFGPRSQ--- 117
DB 61 LRRRGDFVSLQAWTPVVLNGLAAVREALVTRGEDTADRPVPIQILGFGPRSQGLF 120
QY 118 -----GRPRPNGLDOK 129
DB 121 LARYCPANREQRFVSITURNLGLGKKSLEQWTEATYVLCAPADHAGRPFRPNGLDOK 180
QY 130 AVSNVIAISLTCRRFEYDDPRELRLLDLAQGLKEESGFLREVLNAVPLVLLHIALAGKV 189
DB 181 AVSNVIAISLTCRRFEYDDPRELRLLDLAQGLKEESGFLREVLNAVPLVLLHIALAGKV 240
QY 190 LRFQKAFITQLDELLTEHRMTWDPAPPRDLTEAFIAEMKAKGNPSSFNDENRIVVA 249
DB 241 LRSQKAFIAQLDELLTEHRMTWDPAPPRDLTEAFIAEMKAKGNPSSFNDENRIVVA 300
QY 250 DLFSAGMTTSTTLAWGLLMLTHLPDQVRVQOEITDDVIGVRRPEMGDAHMPYTTAVI 309
DB 301 DLFSAGMTTSTTLAWGLLMLTHLPDQVRVQOEITDDVIGVRRPEMGDAHMPYTTAVI 360
QY 310 HEVQRFQDIPVPLGVTHMTSRDIEVQGRIPKGTTLITLSSVLKDEAVMEKPFRRPHEHF 369
DB 361 HEVQRFADIPVPLGVTHMTSRDIEVQGRIPKGTTLITLSSVLKDEAVMEKPFRRPHEHF 420
QY 370 LDAQGHFKVPEAFIPFSAGRRACGLPELARMELFLFTSLLOHFSFVSPTGQPRSHGV 429
DB 421 LDAQGRFKVPEAFIPFSAGRRACGLPELARMELFLFTCLLQRFVSFVPAGQPRSPHGV 480
QY 430 FAFLVTPSPYELCAVPR 446

```

```

DB 481 FAFLVTPSPYELCAVPR 497
|||||
RESULT 4
CDP_PIG STANDARD; PRT; 499 AA.
AC 046658;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D25 (EC 1.14.14.-) (CYP1D25) (Vitamin D(3) 25-
DE hydroxylase).
GN CYP2D25.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-56; 248-272 AND 407-429.
RC TISSUE=Liver;
RX MEDLINE=98086378; PubMed=9425298;
RA Postlind H., Axen E., Bergman T., Wikvall K.;
RT "Cloning, structure, and expression of a cDNA encoding vitamin D3 25-
RT hydroxylase.";
RL Biochem. Biophys. Res. Commun. 241:491-497(1997).
RN [2]
RP SEQUENCE OF 1-16.
RC TISSUE=Liver;
RX MEDLINE=93075023; PubMed=1445236;
RA Axen E., Bergman T., Wikvall K.;
RT "Purification and characterization of a vitamin D3 25-hydroxylase
RT from pig liver microsomes.";
RL Biochem. J. 287:725-731(1992).
CC -!- FUNCTION: CATALYZES THE FIRST STEP IN THE METABOLIC ACTIVATION OF
CC VITAMIN D(3) INTO 1-ALPHA,25-DIHYDROXYVITAMIN D(3), ITS ACTIVE,
CC HORMONAL FORM.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- TISSUE SPECIFICITY: Found in liver and kidney.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y16417; CAA76205.1; -.
CC PIR; JC5819; JC5819.
CC HSP; P00179; 1DT6.
CC InterPro; IPR001128; Cytochrome P450.
CC InterPro; IPR008069; EP450_CYP2D.
CC Pfam; PF00067; p450; 1.
CC PRINTS; PR01686; EP450CYP2D.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT INIT MET 0
FT METAL 445 445 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 499 AA; 56380 MW; 31C878B580E61919 CRC64;
Query Match 74.5%; Score 1736.5; DB 1; Length 499;
Best Local Similarity 69.7%; Pred. No. 1.6e-115;
Matches 341; Conservative 39; Mismatches 58; Indels 51; Gaps 1;
QY 9 LAVIVAIPELLVLDLHRRQRWAARYSPGLPLGGLNLLHVDVFONTPCYDQRRRFGDV 68
DB 11 LALAMVIFELLVLDLHRRSRWAPRYPPGMPGLGGLNLLQVNFQDPRLSFIQLRRRFGDV 70
QY 69 FSLQLAWTPVVVNLGLAAVREALVTHGEDTADRPVPIQILGFGPRSQ----- 117

```

```
Db 71 FSLQIWRVWVNLGLAAVREALVSHSHSETSDRPPVFILEHLGYPGRSEGVILARYGKAW 130
Qy 118 -----GRFRPNGLLDKAVSNVIA 137
Db 131 RQRPSVSTLNFGLGKSLSEWVTOEASCLCAAFADQAGFPFSPNLLNKA VSNVIA 190
Qy 138 LTCGRREYDDPRFLRLDLAEGLEKESGFLREVNLAVPVLLHIPALAGKVLRFQKAPL 197
Db 191 LTFARRFEYNDPRMLKLLDLLEGLKEEVGLMRQVLEAMPVLRHIPGLCAKLPFRQKAPL 250
Qy 198 TOLDELLTEHRTWDPAPQPRDLTEAFLEAMEKAKGNPSSFNNDENLRIVVADLPSAGNV 257
Db 251 VMIDELITEHKTRDLAQPRLTDAFLDEMEKAKGNPSSFNNDENLRVVAHLFSAGMI 310
Qy 258 TTSTTLAWGLMLLHPDVQRRVQOEIDVIGVRRPEMGDQAHMPYTTAVIHEVQRF 317
Db 311 TTSTTLAWGLMLLHPDVQRRVQOEIDVIGVRRPEMGDQAHMPYTTAVIHEVQRF 370
Qy 318 IVPLGVHMTSDIEVQGPRIKPGTTLITNLSSVLKDEAVWEKPFRRFHPHFELDAQGHV 377
Db 371 IVPLGVHMTSDIEVQGPRIKPGTTLITNLSSVLKDEAVWEKPFRRFHPHFELDAQGRFT 430
Qy 378 KPEAFLEPSAGRACGEPLARMELFFLFTSLLOHFSFVSPTQPRPSHHGVFAFLVTPS 437
Db 431 KQEAEMPPFSAGRRSCLGEPLARMELFFLFTSLLOHFSFVSPTQPRPSHHGVFAFLVTPS 490
Qy 438 PYELCAVPR 446
Db 491 PYQLCAVPR 499
```

RESULT 5

```
CPDE BOVIN
ID CPDE BOVIN STANDARD; PRT; 487 AA.
AC Q01361;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D14 (EC 1.14.14.1) (CYP1D14) (Fragment).
GN CYP2D14.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93011103; PubMed=1396678;
RA Tsuneoka Y., Matsuo Y., Higuchi R., Ichikawa Y.;
RT "Characterization of the cytochrome P-4501D subfamily in bovine
RT liver. Nucleotide sequences and microheterogeneity.";
RL Eur. J. Biochem. 208:739-746(1992).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC or send an email to license@isb-sib.ch).
DR EMBL; X68013; CAA48149.1; -.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT NON_TER 1
FT METAL 433 433 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 487 AA; 54687 MW; D61CFE3B8ADE19E7 CRC64;
Query Match 74.4%; Score 1734.5; DB 1; Length 487;
Best Local Similarity 69.4%; Pred. No. 2.1e-115;
Matches 338; Conservative 40; Mismatches 58; Indels 51; Gaps 1;
Qy 11 VIVATFLLLVDLMHRRQRWAARYSPGLPLGGLGNLLHVDFTQTPYCFDQLRRRFGDVS 70
Db 1 VALLIFLLLLDLMHRRSRWAPYPGPTPLVLGNLLQVDFEDPRPSFNQLRRRFGNVS 60
Qy 71 LQLAWTPVVLNGLAAVREALVTHGEDTADRPVPITQILGEGPRSQ----- 117
Db 61 LQQWTPVVLNGLAAVREALVYRSQDTADRPVAVYERLGYGPRAEGLVILARYGDAMRE 120
Qy 118 -----GRFRPNGLLDKAVSNVIA 139
Db 121 QRRFSLTTRNFGLGKSLSEWVTEASCLCAAFADQAGRPSPNLLNKA VSNVIA 180
Qy 140 CGRREYDDPRFLRLDLAEGLEKESGFLREVNLAVPVLLHIPALAGKVLRFQKAPL 199
Db 181 FGRREYNDPRILKLLDTEGLKEFNLVKVEAVPVLLSIPGLAARVFAQAFMAL 240
Qy 200 LDELTEHRTWDPAPQPRDLTEAFLEAMEKAKGNPSSFNNDENLRIVVADLPSAGMVT 259
Db 241 IDGLAEQKMTQDPTQPPRLHLDLDAFLDEMEKAKGNPSSFNNDENLRVVAHLFSAGMVT 300
Qy 260 STTLAWGLMLLHPDVQRRVQOEIDVIGVRRPEMGDQAHMPYTTAVIHEVQRF 319
Db 301 STTLAWGLMLLHPDVQRRVQOEIDVIGVRRPEMGDQAHMPYTTAVIHEVQRF 360
Qy 320 PLGVHMTSDIEVQGPRIKPGTTLITNLSSVLKDEAVWEKPFRRFHPHFELDAQGHV 379
Db 361 PLGLPHMTSDIEVQGFHPIKGTTLITNLSSVLKDEAVWEKPFRRFHPHFELDAQGHV 420
Qy 380 EAFLEPSAGRACGEPLARMELFFLFTSLLOHFSFVSPTQPRPSHHGVFAFLVTPSPY 439
Db 421 EAFLEPSAGRACGEPLARMELFFLFTSLLOHFSFVSPTQPRPSHHGVFAFLVTPSPY 480
Qy 440 ELCAVPR 446
Db 481 QLCAVPR 487
RESULT 6
CPD4_RAT
ID CPD4_RAT STANDARD; PRT; 500 AA.
AC P13108; O35107;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D4 (EC 1.14.14.1) (CYP1D4) (P450-DB4) (P450-CMF3)
DE (Debrisoquine 4-hydroxylase).
GN CYP2D4 OR CYP2D-4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
```



```
DR EMBL; U48219; AAC52882.1; -.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 446 446 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 500 AA; 56683 MW; 9848A8BE5ABA09C5 CRC64;

Query Match 73.3%; Score 1708.5; DB 1; Length 500;
Best Local Similarity 68.9%; Pred. No. 1.5e-113;
Matches 337; Conservative 33; Mismatches 71; Indels 51; Gaps 1;

QY 6 LVPLAVIAIFLLVLDLHRRORWAARYSPGLPLGCLGNLLHVDVONTYPYCFDQLRRP 65
DB 9 LPIAIFTIFLLVLDLHRRORWRTSRYPGPVFWVLGNLLQIDFQNNPAGFQKLRCP 68
QY 66 GDVFSQLAWTPVWVLNGLAAREALVTHGDETADRPVPVITQILGFGPRSQ----- 117
DB 69 GDLFSLQAFESVVLNGLPALREALVKYSEDADRPPLHFNDQSGFGPRSQGVVLARYG 128
QY 118 -----GRFPRNGLLDKAVSNV 134
DB 129 PAWROQRFSVSTRFHGLGKSLQWVTEARCLCAAFADHSGFPFSPNTLLDKAVCNV 188
QY 135 IASLTGRRFEYDDPFRLLDLAQSLKEESFLREVLNAVVLHIIPALACKVLRFOK 194
DB 189 IASLLFACFEYNDPFRLLDLQDLTEESGFLPMLLNVPMLLHIIPGLGKFSGKK 248
QY 195 AFITQDDELLTEHRTMTDPAQPRDITAEFLAEMEKAKGNPSSFNDELRIYVADLFA 254
DB 249 AFVAMDELLETKHTWDPAQPRDITDAFLAEVEKAKGNPSSFNDELRIYVADLFA 308
QY 255 GMYTTSTLAWGLLMLHPDVORRVOQEIIDDVIGVRRPEMGDAQHPYTTAVIHEVOR 314
DB 309 GMYTTSTLWALLFMRPDVQCRVQOEIDEVIGVRRPEMADQARMPFTNAVIEVQR 368
QY 315 FGDIVPLGVTHMSTRDIEVQGFRIKGTTLITNLSSVLKDEAWKXPFPHFHFDAOG 374
DB 369 FADILPLGVPHKTSRDIEVQGFRIKGTTLITNLSSVLKDETWKXPFPHFHFDAOG 428
QY 375 HFVKPFAFLPFSAGRAACGLCEPLARMELFLFTSLQHFSFSVPTQOPRPSHHGVFAFLV 434
DB 429 NFVKHFAFPFSAGRAACGLCEPLARMELFLFTCLLQRFSSVPAGQPRPSNYGVFGALT 488
QY 435 TSPSYELCAVPR 446
DB 489 TPRPYQLCASPR 500

RESULT 8
CPDF CANFA
ID CPDF CANFA STANDARD; PRT; 499 AA.
AC Q29473; Q02859;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D15 (EC 1.14.14.1) (CYP11D15) (P450 DUT2).
GN CYP2D15.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC PROSITE=Liver;
RX MEDLINE=95305574; PubMed=7786018;
RA Sakamoto K., Kirita S., Baba T., Nakamura Y., Yamazoe Y., Kato R.,
RA Takanaka A., Matsubara T.;
```

```
Db 245 AQAIAITLTNEMIQSHKRTDTPQPRHLIDAFVDEIEKAKGNPKTSFNEENLCNWTSDL 304
Qy 252 FSAGVMTTSTTLAWGLLLMLHPDVQRRVQOQIDVIGVRRPENGDAQHMPYTTAVIHE 311
Db 305 FIAGMVSTSTLTWALLMLHPDVQRRVQOQIDVIGVRRPENGDAQHMPYTTAVIHE 364
Qy 312 VQREGDVIPLGVTHMTSDRIEVOGFRIPKGTTLITNLSSVLKDEAVWEKPRFHFHEFLD 371
Db 365 VQREGDVIPLGVTHMTSDRIEVOGFRIPKGTTLITNLSSVLKDEAVWEKPRFHFHEFLD 424
Qy 372 AQGHEVKEAFPLFSSAGRRACGLGEPLARMELFFFTSLQLHFSFVSVPQTGQPRPSHHGVFA 431
Db 425 AQGHEVKEAFPLFSSAGRRACGLGEPLARMELFFFTSLQLHFSFVSVPQTGQPRPSHHGVFA 484
Qy 432 FLVTPSPYELCAVPR 446
Db 485 FLKVPAPQLCQVEPR 499

RESULT 9
CPDK MESAU
ID CPDK MESAU STANDARD; PRT; 500 AA.
AC Q9OYG5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D20 (EC 1.14.14.-) (CYP1D20).
GN CYP2D20.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20533996; PubMed=11083025;
RA Oka T., Fukuhara M., Ushio F., Kurose K.;
RT "Molecular cloning and characterization of three novel cytochrome
RT P450 2D isoforms, CYP2D20, CYP2D27, and CYP2D28 in the Syrian hamster
RT (Mesocricetus auratus).";
RL Comp. Biochem. Physiol. 127C:143-152(2000).
RC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By
CC similarity).
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB031864; BAA99313.1; -.
DR HSSP; P00179; 1D76.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; p450.1.
DR PRINTS; PR01686; EP4501CYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; MONOOXYGENASE; Electron transport; Membrane; Heme;
KW Oxidoreductase; Monooxygenase;
FT METAS. 446 446
FT METAL. 446 446
SQ SEQUENCE 500 AA; 56503 MW; 9948102706C50144 CRC64;
```

Query Match 71.0%; Score 1653.5; DB 1; Length 500;
Best Local Similarity 66.5%; Pred. No. 1.2e-109;
Matches 324; Conservative 33; Mismatches 79; Indels 51; Gaps 1;

```
Qy 11 VIVAFILLVLMHRRQRAARYSPGLPLGLGNLHVDQNTPCYCFDQLRRRREGDVFS 70
Db 14 IFTALFLVLVLMHRRKFRARYPPGPMPLGLGNLQVDFENMPEYLYKFKQORYGDVFS 73
Qy 71 LQLAWTPVVVINGLAAVREALVTHGEDTADRPVPVITQILGFGPRSQ----- 117
Db 74 LQMAWKPVVVINGLKAVERVLVNCGEDTADRPVPVIFNHLGYRPSQGVVFARYGPQWRE 133
Qy 118 -----GRPFRPNGLLDKAVSNVIALIT 139
Db 134 QRRFSVSTMRDVGKKSLQEWVTEEAGHLCDAFTQEAQHPFNPTITLKNKSVCNVSSLI 193
Qy 140 CGRFEYDDPFLRLDLAQGLKEESGFLREVLNAVPLLHHPALACKVLRFQKALITQ 199
Db 194 YAHRPDYEDPFFNKLLKTLQESFGEDSGFIAEVLNAVPLLRIPGLPKAPKLTAFMDS 253
Qy 200 LDELTEHRMTWDPAPPRDLTEAFLEMEKAKGNPSSFNDENLRIVVADLFSAGMVT 259
Db 254 LYKMLIEHKTWDPAPPRGLTDAFLAEVAKGRPESSFNDENLHVVVADLFIAGMVT 313
Qy 260 STTLAWGLLLMLHPDVQRRVQOQIDVIGVRRPENGDAQHMPYTTAVIHEVQRFQDIV 319
Db 314 STTLAWGLLLMLHPDVQRRVQOQIDVIGVRRPENGDAQHMPYTTAVIHEVQRFQDIV 373
Qy 320 PLGVTHMTSDRIEVOGFRIPKGTTLITNLSSVLKDEAVWEKPRFHFHEFLDAQHFKVP 379
Db 374 PNVPHMTSDRIEVOGFRIPKGTTLITNLSSVLKDEAVWEKPRFHFHEFLDAQHFKVP 433
Qy 380 EAFPLFSSAGRRACGLGEPLARMELFFFTSLQLHFSFVSVPQTGQPRPSHHGVFAFLVTPSPY 439
Db 434 EAFPLFSSAGRRACGLGEPLARMELFFFTSLQLHFSFVSVPQTGQPRPSHHGVFAFLVTPSPY 493
Qy 440 ELCAVPR 446
Db 494 ELCAVPR 500

RESULT 10
CPDK MESAU
ID CPDK MESAU STANDARD; PRT; 500 AA.
AC Q9OYG6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D27 (EC 1.14.14.-) (CYP1D27).
GN CYP2D27.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=20533996; PubMed=11083025;
RA Oka T., Fukuhara M., Ushio F., Kurose K.;
RT "Molecular cloning and characterization of three novel cytochrome
RT P450 2D isoforms, CYP2D20, CYP2D27, and CYP2D28 in the Syrian hamster
RT (Mesocricetus auratus).";
RL Comp. Biochem. Physiol. 127C:143-152(2000).
CC -!- FUNCTION: Has bufuralol 1'-hydroxylase and debrisoquine 4-
CC hydroxylase activities.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By
CC similarity).
CC -!- TISSUE SPECIFICITY: Expressed in liver, but not in kidney, small
CC intestine, and brain.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```


DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 446 446 IRON (HEME AXIAL LIGAND).
FT CONFLICT 117 117 N -> D (IN REF. 3).
FT CONFLICT 346 346 R -> L (IN REF. 1 AND 2).
FT CONFLICT 358 358 F -> L (IN REF. 1 AND 2).
FT CONFLICT 407 407 K -> E (IN REF. 1 AND 2).
SQ SEQUENCE 500 AA; 56683 MW; 23B99250734C2215 CRC64;

Query Match 70.4%; Score 1640; DB 1; Length 500;
Best Local Similarity 64.0%; Pred. No. 1.1e-108;
Matches 320; Conservative 47; Mismatches 79; Indels 54; Gaps 2;

QY 1 MGL---EALVPLAVIAIFLLVDMHRRQRAARYSPGFLPLGLGNLLHVDFTQNYTC 57
Db 1 MGLIGDGLLWAVIFTAIFLLVDMHRRQRAARYSPGFLPLGLGNLLHVDFTQNYTC 60

QY 58 FQLRRRFGDVSLOLAWTPVVLNGLAAVREALVTHGDDTADRPVPITQILGFGPRSQ 117
Db 61 LYKLRSRYGDVFSLOLAWTPVVLNGLAAVREALVTHGDDTADRPVPITQILGFGPRSQ 120

QY 118 G-----RPRPNGL 126
Db 121 GVVLAPYGEWEQRFRFVSVDLDRFGVGKSLQEWTEAGHLCDTFAKEAHPNPSIL 180

QY 127 LDKAVSNVIASLTGRRFFDYDPRFLRLDLAQEGLKBESGFLREVLNAVPVLLHLPALA 186
Db 181 LSKAVSNVIASLVYARRFEYDFEFNMLKTLKESGEGDTGPMVAELNAPILLQIPGLP 240

QY 187 GKVLRFQKFLQDLLELTHRMWDPAQPPDLTEAFLEAMEKAKGNPESFNDENLRI 246
Db 241 GKVPFLKLSFIALVDKMLIEHKSKDPAQPPDMTDAFLAEMQKAKGNPESFNDENLRL 300

QY 247 VVADLFAGMVTSTTTLAWGLLMLHPDVQRRVQEDIDVIGVRRPDMGQAHMPYTT 306
Db 301 VVIDLFAGMVTSTTTLAWGLLMLHPDVQRRVQEDIDVIGVRRPDMGQAHMPYTT 360

QY 307 AVIHEVRFGDIPVLGVTHMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPRFRHP 366
Db 361 AVIHEVRFGDIPVLGVTHMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPRFRHP 420

QY 367 EHLDAQGHVPEAFPLPSAGRRACLGEPFLARMELFLFTSLLOHFSFVPTGQPRSH 426
Db 421 EHLDAQGHVPEAFPLPSAGRRACLGEPFLARMELFLFTSLLOHFSFVPTGQPRSH 480

QY 427 HGVEAFVLPSPYELCAVER 446
Db 481 HGVEAFVLPSPYELCAVAR 500

RESULT 12

CDP3 RAT
ID CDP3 RAT STANDARD; PRT; 500 AA.
AC P12938; O35106;
DT 01-OCT-1989 (Rel. 12, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D3 (EC 1.14.14.1) (CYP1D3) (P450-DB3) (Debrisoquine 4-hydroxylase).
GN CYP2D3 OR CYP2D-3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RX MEDLINE=90057430; PubMed=2819073;
RA Matsunaga T., Zanger U.M., Hardwick J.P., Gelboin H.V., Meyer U.A., Gonzalez F.J.;

RT "The CYP2D gene subfamily: analysis of the molecular basis of the
RT debrisoquine 4-hydroxylase deficiency in DA rats.";
RL Biochemistry 28:7349-7355(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=90189185; PubMed=2107330;
RA Matsunaga E., Umeno M., Gonzalez F.J.;
RT "The rat P450 IID subfamily: complete sequences of four closely
RT linked genes and evidence that gene conversions maintained sequence
RT homogeneity at the heme-binding region of the cytochrome P450 active
RT site.";
RL J. Mol. Evol. 30:155-169(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=9809365; PubMed=9434752;
RA Wan J., Inaoka S., Chow T., Hiroi T., Yabueki Y., Funae Y.;
RT "Expression of four rat CYP2D isoforms in Saccharomyces cerevisiae
RT and their catalytic specificity".
RL Arch. Biochem. Biophys. 348:383-390(1997).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J02868; AAA41002.1; -.
DR EMBL; X52028; CAA36270.1; -.
DR EMBL; AB008424; BAA23124.1; -.
DR PIR; S16872; S16872.
DR HSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 446 446 IRON (HEME AXIAL LIGAND).
FT CONFLICT 125 126 AP -> CT (IN REF. 1 AND 2).
SQ SEQUENCE 500 AA; 56641 MW; C54727C2C00F73F6 CRC64;

Query Match 70.3%; Score 1638.5; DB 1; Length 500;
Best Local Similarity 64.8%; Pred. No. 1.3e-108;
Matches 319; Conservative 43; Mismatches 79; Indels 51; Gaps 1;

QY 6 LVPLAVIAIFLLVDMHRRQRAARYSPGFLPLGLGNLLHVDFTQNYTCFQDLRRRF 65
Db 9 LWPMAIFVIFLLVDMHRRQRAARYSPGFLPLGLGNLLHVDFTQNYTCFQDLRRRF 68

QY 66 GDVFSLOLAWTPVVLNGLAAVREALVTHGDDTADRPVPITQILGFGPRSQ----- 117
Db 69 GDVFSLOLAWTPVVLNGLAAVREALVTHGDDTADRPVPITQILGFGPRSQ----- 128

QY 118 -----GRFPFRPNGLDKAVSNV 134
Db 118 -----GRFPFRPNGLDKAVSNV 134

```

Db 129 PEWREORRFSVTLNRFNGVKKSLQWVTDASHLCDALTAEAGRPLDPYTLNKAQVNV 188
QY 135 IASLTGCRREFYDDPRFLRLDLAQGLKEESGFLREVNLNAVPLHHPALAGKVLRFQX 194
Db 189 IASLIYARFDYGDPDFIKVILKESMGEOGLFEVLNMPVLLRIPLGLADKVPFGQX 248
QY 195 AFLTQDELLELTHMTWMDPAQPPRLDTEAFLEAMEKAKGNPSSFNDELRIVADLFS 254
Db 249 TELTMDVNLVTEHKKTWDPQPPRLDTEAFLEAMEKAKGNPSSFNDELRIVADLFS 308
QY 255 GWTVTSTLAWGLLMLHPDQVRVQOEIDVIGOVREPMCDQAHMPTTAVIHEVOR 314
Db 309 GWTVTSTLAWGLLMLHPDQVRVQOEIDVIGOVREPMCDQAHMPTTAVIHEVOR 368
QY 315 FGDIVPLGVTHMTSRDIEYVQGFRIKPGTTLINLSSVLKDEAVWEKPFHFHEHFLDAQ 374
Db 369 FADIVPMNLPHKTSRDIEYVQGFELIPKGTTLINLSSVLKDEAVWEKPFHFHEHFLDAQ 428
QY 375 HFVKPEAFIPFSGRACIAGEPLARMELFFTLQHFSSVPTQPPRSHHGVPFVLV 434
Db 429 NFVKEAFPFSGRACIAGEPLARMELFFTLQHFSSVPTQPPRSHHGVPFVLV 488
QY 435 TPSPYELCAVPR 446
Db 489 SPSPYOLCAVPR 500

RESULT 13
CPDL_RAT
ID CPDL_RAT STANDARD; PRT; 504 AA.
AC P10633; O35105;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2B1 (EC 1.14.14.1) (CYP2B1) (P450-CMP1A)
DE (P450-UT-7) (Debrisoquine 4-hydroxylase).
GN CYP2D1 OR CYP2D-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=101116;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=87217961; PubMed=3582092;
RA Gonzalez F.J.; Matsunaga T.; Nagata K.; Meyer U.A.; Nebert D.W.;
RA Pastewka J.; Kozak C.A.; Gillette J.; Gelboin H.V.; Hardwick J.P.;
RT "Debrisoquine 4-hydroxylase: characterization of a new P450 gene;
RT the DA rat polymorphism.";
RL DNA 6:149-161(1987).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=90057430; PubMed=2819073;
RA Matsunaga T.; Zanger U.M.; Hardwick J.P.; Gelboin H.V.; Meyer U.A.;
RA Gonzalez F.J.;
RT "The CYP2D gene subfamily: analysis of the molecular basis of the
RT debrisoquine 4-hydroxylase deficiency in DA rats.";
RL Biochemistry 28:7349-7355(1989).
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=89050091; PubMed=3190674;
RA Ishida N.; Tawaragi Y.; Inuzuka C.; Sugita O.; Kubota I.;
RA Nakazato H.; Noguchi T.; Sasa S.;
RT "Four species of cDNAs for cytochrome P450 isozymes immunorelated to
RT P450C-M/F encode for members of P450IID subfamily, increasing the
RT number of members within the subfamily.";
RL Biochem. Biophys. Res. Commun. 156:681-688(1988).
RN [4]
RN SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=9809365; PubMed=9434752;
RA Wan J.; Imaoka S.; Chow T.; Hiroi T.; Yabusaki Y.; Funae Y.;
RT "Expression of four rat CYP2D isoforms in Saccharomyces cerevisiae

```

```

RT and their catalytic specificity.";
RL Arch. Biochem. Biophys. 348:383-390(1997).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M16654; AAA41054.1; -
CC EMBL; J02867; AAA41001.1; -
CC EMBL; M22328; AAA41043.1; -
CC EMBL; AB008422; BAA23122.1; -
CC PIR; A26822; A26822.
CC HSP; P00179; I076.
CC InterPro; IPR001128; Cytochrome_P450.
CC InterPro; IPR008069; EP450_CYP2D.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR01686; EP450ICYP2D.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 446 446 IRON (HEME AXIAL LIGAND).
FT CONFLICT 123 124 IL -> VF (IN REF. 3).
FT CONFLICT 173 173 Q -> R (IN REF. 3).
FT CONFLICT 380 380 F -> I (IN REF. 3 AND 4).
SQ SEQUENCE 504 AA; 57175 MW; 2F9AD87B4EF327DC CRC64;

Query Match 69.3%; Score 1614.5; DB 1; Length 504;
Best Local Similarity 64.6%; Pred. No. 6.8e-107;
Matches 316; Conservative 38; Mismatches 84; Indels 51; Gaps 1;

QY 9 LAVIVAIFLLVLDLHRRQRWAARYSPGPLPLPGLGNLHVDFTQPTCYDQLRRFGDV 68
Db 12 MAIFTVFIFLLVLDLHRRHRMTSRYPGVPVPLGLNLLQVLDLSNNPYSLYKLQHYGDV 71
QY 69 FSLQAWTPVVLNGLAVREALVTHGEDTADPPVPIITQILGFGPRSQ----- 117
Db 72 FSLQGWKPMVIVNRLKAVQEVLTGHTDADPPVPIFKICLGVKRSQGVILASVGPWE 131
QY 118 -----GRPFRLNGLDKAVSNVIAS 137
Db 132 REQRFRFSVSTLRTFGMKKSLSEWTKKAGHLCDAFTAQAGQSQINPKAMLNKALCNVIAS 191
QY 138 LTCGRREFYDDPRFLRLDLAQGLKEESGFLREVNLNAVPLHHPALAGKVLRFQX 197
Db 192 LIFARRFEYEDPYLRIMVXKLVESLTVSGFPEVLTNPALLRIPGLADKVPFGQKTFM 251
QY 198 TOLDLLELTHMTWMDPAQPPRLDTEAFLEAMEKAKGNPSSFNDELRIVADLFSAGMV 257
Db 252 ALLDNLAEARTTWDPAQPPRLDTEAFLEAMEKAKGNPSSFNDELRIVADLFSAGMV 311
QY 258 TSTTLAWGLLMLHPDQVRVQOEIDVIGOVREPMCDQAHMPTTAVIHEVORFGD 317
Db 312 TTAATLITWALLMLLIPYDQVRVQOEIDVIGOVREPMCDQAHMPTTAVIHEVORFGD 371
QY 318 IVPLGVTHMTSRDIEYVQGFRIKPGTTLINLSSVLKDEAVWEKPFHFHEHFLDAQHFV 377

```

Db 372 IAPLNLPRFTSDIEVQDFVIPKGTTLIINLSVLKDETWEKPHRPHFPEHFLDAQGNFV 431
 Qy 378 KPEAFIPSSAGRACLGEPFLARMELFLPFTSLQHFSSVPTGQPRPSHHGVFAFLVTPS 437
 Db 432 KHEAFMPFSAGRACLGEPFLARMELFLPFTCLQRFSPVPTGQPRPSHHGVFAFLV 491
 Qy 438 PYELCAVPR 446
 Db 492 FYQLCAVPR 500

RESULT 14
 CPD5_RAT
 ID CPD5_RAT STANDARD; PRT; 504 AA.
 AC P12379.
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 2D5 (EC 1.14.14.1) (CYP11D5) (P450-DB5) (P450-CMF1B)
 DE (Debrisoquine 4-hydroxylase).
 GN CYP2D5 OR CYP2D-5.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=90189185; PubMed=2107330;
 RA Matsunaga E., Umeno M., Gonzalez F.J.;
 RT "The rat P450 IID subfamily: complete sequences of four closely
 linked genes and evidence that gene conversions maintained sequence
 homogeneity at the heme-binding region of the cytochrome P450 active
 site.";
 RL J. Mol. Evol. 30:155-169(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90057430; PubMed=2819073;
 RA Matsunaga T., Zanger U.W., Hardwick J.P., Gelboin H.V., Meyer U.A.,
 Gonzalez F.J.;
 RT "The CYP2D gene subfamily: analysis of the molecular basis of the
 debrisoquine 4-hydroxylase deficiency in DA rats.";
 RL Biochemistry 28:7349-7355(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89366685; PubMed=2771656;
 RA Ishida N., Inuzuka C., Tawaragi Y., Sugita O., Nakazato H.,
 Noguchi T., Sassa S., Kappas A.;
 RT "Cytochrome P450CMF cDNA: nucleotide sequence of P450CMF1b.";
 RL Nucleic Acids Res. 17:6407-6407(1989).
 RN [4]
 RP SEQUENCE OF 18-504 FROM N.A.
 RX MEDLINE=89050091; PubMed=3190674;
 RA Ishida N., Tawaragi Y., Inuzuka C., Sugita O., Kubota I.,
 Nakazato H., Noguchi T., Sassa S.;
 RT "Four species of cDNAs for cytochrome P450 isozymes immunorelated to
 P450-M/F encode for members of P450IID subfamily, increasing the
 number of members within the subfamily.";
 RL Biochem. Biophys. Res. Commun. 156:681-688(1988).
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 monooxygenases. In liver microsomes, this enzyme is involved in an
 NADPH-dependent electron transport pathway. It oxidizes a variety
 of structurally unrelated compounds, including steroids, fatty
 acids, and xenobiotics.
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -!- INDUCTION: P450 can be induced to high levels in liver and other
 tissues by various foreign compounds, including drugs, pesticides,
 and carcinogens.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M25143; AAA41034.1; -;
 CC EMBL; X52030; CAA36272.1; -;
 CC EMBL; J02869; AAA41003.1; -;
 CC EMBL; M22329; AAA41045.1; -;
 CC PIR; S09611; O4RTD5.
 CC HSP; P00179; 1DT6.
 CC InterPro; IPR001128; Cytochrome P450.
 CC InterPro; IPR008069; EF450_CYP2D.
 CC Pfam; PF00067; P450; 1.
 CC PRINTS; PR01686; EP450ICYP2D.
 CC PRINTS; PR00385; P450.
 CC PROSITE; PS00086; CYTOCHROME P450; 1.
 CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT METAL 446 446 IRON (HEME AXIAL LIGAND).
 SQ SEQUENCE 504 AA; 57076 MW; F308EE06F2605DFE CRC64;
 Query Match 69.3%; Score 1614.5; DB 1; Length 504;
 Best Local Similarity 64.6%; Pred. No. 6.8e-107;
 Matches 318; Conservative 38; Mismatches 85; Indels 51; Gaps 1;
 Qy 6 LVPLAVIVAIPELLLVLMHRRQRWAARYSPGLPLPGLGNLLHVDFOQNTYCFDQRRRF 65
 Db 9 LWPMAITFTVIFILLVDMHRRQRWTSRYPPGVPWVPLGNLLQVDPSNWPYSMYKLQHY 68
 Qy 66 GDVFSQLQAWTPVVVINGLAARVREALVTHGDDTADRPVPITQILGQPRSQ----- 117
 Db 69 GDVFSLOWGKWPVIVNRLKAVQEVLVTHGDDTADRPVPVIFKCLGVKPRSQGVVFSYG 128
 Qy 118 -----GRPPFNGLLDKAVSNV 134
 Db 129 PEWREQRFRSVSTLRTFGMGKKSLEEMWYKKAHGLCDAFTAQNGRSINPKAMLNKALCNV 188
 Qy 135 IASLTGCRFRFYDDPRFLRLDLAQEGLKESGFLREVINAVPVLLHLPALAGKVLRFQK 194
 Db 189 IASLIFARRFEYEDPYLLIRMLTLVEESLIEVSGFIPEVLTFFALLRPLGLADKVFQOK 248
 Qy 195 AFLTQDELLETHRWTDPAQPRDLTEAFIAEMEKAKGNPSSFNDENLRIVVADLPSA 254
 Db 249 TFMAFLDNLAAENRTTWDPAPPRNLDAFLAEVEKAKGNPSSFNDENLRIVVADLPSA 308
 Qy 255 GMVTTSTTLAWGLLMLHPDVQRRVQOEIDVTIGVRRPEMGDQAHMPYTTAVIHEVQR 314
 Db 309 GMVTTATTLTALLMLIPDVQRRVQOEIDVTIGVRRPEMGDQAHMPYTTAVIHEVQR 368
 Qy 315 FGDIVPLGVTHMTRSDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPRFRPHFPEHFLDAQG 374
 Db 369 FGDIAPLNLPRTITSCDIEVQDFVIPKGTTLIINLSVLKDETWEKPLRFPHEHFLDAQG 428
 Qy 375 HFVKEAPFLPSSAGRACLGEPFLARMELFLPFTSLQHFSSVPTGQPRPSHHGVFAFLV 434
 Db 429 NFVKEAFMPFSAGRACLGEPFLARMELFLPFTCLQHFSSVPTGQPRPSHHGVFAFLV 488
 Qy 435 TPSPYELCAVPR 446
 Db 489 APLPYQLCAVPR 500

RESULT 15
 CPDG_CAVPO
 ID CPDG_CAVPO STANDARD; PRT; 500 AA.
 AC Q64403; O54866;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 2D16 (EC 1.14.14.1) (CYP11D16).

GN CYP2D16.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1] SEQUENCE FROM N.A., AND SEQUENCE OF 1-38.
RP STRAIN=13; TISSUE=Adrenal cortex;
RC MEDLINE=95251703; PubMed=7733969;
RX Jiang Q., Voigt J.M., Colby H.D.;
RA "Molecular cloning and sequencing of a guinea pig cytochrome P4502D
RT (CYP2D16): high level expression in adrenal microsomes";
RL Biochem. Biophys. Res. Commun. 209:1149-1156(1995).
RN [2] SEQUENCE FROM N.A.
RP STRAIN=Hartley white; TISSUE=Adrenal gland;
RA Sun Y., Voigt J.M., Pierce J.C., Colby H.D.;
RT "The gene sequence of a xenobiotic metabolism-related cytochrome P450
RT isozyme (CYP2D16) in guinea pig adrenal gland";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE INNER ZONE OF
CC THE ADRENAL CORTEX.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U21486; AAC68479.1; --
CC EMBL; AF020345; AAB94568.1; --
CC FIR; JC4153; JC4153.
CC HSP; P00179; I076.
CC InterPro; IPR001128; Cytochrome P450.
CC InterPro; IPR008069; EP450_CYP2D.
CC Pfam; PF00067; P450; 1
CC PRINTS; PR01686; EP450ICYP2D.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT DOMAIN 81 84 POLY-VAL.
FT METAL 446 446 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT CONFLICT 123 123 I -> V (IN REF. 2).
FT CONFLICT 127 127 Y -> N (IN REF. 2).
FT CONFLICT 148 148 G -> R (IN REF. 2).
SQ SEQUENCE 500 AA; 55800 MW; 2429247E49BF6B24 CRC64;

Query Match 69.1%; Score 1611; DB 1; Length 500;
Best Local Similarity 63.4%; Pred. No. 1.2e-106;
Matches 317; Conservative 44; Mismatches 85; Indels 54; Gaps 2;

QY 1 MGL---EALVPLAVIAIALLLDLMHRRORWAARYSGPLPLGLGNLLHVDVFNQTPYC 57
DB 1 MGLLTGDAIFSVAVAIAIFLLLDLMHRRORWAARYPPGVPVPGVGLNLLQVDNFNMAYS 60

QY 58 FQGLRRRFGDVSQQLATPVVVVNLGLAAVREALVTHGEDTADRPVPVITQILGFGRSQ 117
DB 61 CDKLRHQFGDVSQQLATPVVVVNLGLAAVREALVTHGEDTADRPVPVITQILGFGRSQ 120

QY 118 -----GRFPRNGL 126
|:|:|:|

Db 121 GVIGAYGPAMREQRFRFSVSSLRNFGLGKKSLEQWVTEBAACLCAAFTNHAGQPFCKAL 180
QY 127 LDKAVSNVIASTCGRRFEYDDPRFLRLLDLAQEGLEKEESGFLREVLNAVVPVLLHHPALA 186
Db 181 LNKAVCNVSSLIYARRFDYDDPMVLRLLLEFLEETLRNSSSLKIQVNLNIPILLLRIPCVA 240
QY 187 GKVLRFQKAFLTQDLLELTHEHMTWDPAOPPRDLTEAFLEAEKAKGNPESSEFNENLRI 246
Db 241 AKVLSAQRSFIALNDKLLAEHNTGWAPDQPPRDLTDAFLTEHKAQGNSESSFNENLRL 300
QY 247 VVADLFSAGMVTSTTTLAWGLLLMLHPDVQRRVQOEIDDDVIGOVRRPBGDGAHMPYTT 306
Db 301 LVSDLFAGMVTSTVTLSSWALLMLHPDVQRRVQOEIDDEVIGQVRCPEMADQAHPFFTN 360
QY 307 AVIHEVQRFQDGVPLGVTHMTSRDIEVQGFRIIPKGTTLITNLSSVLKDBAVWEKPRFRHP 366
Db 361 AVIHEVQRFADIVPMGVPHMTSRDTEVQGFLLIPKGTMLFTNLSSVLKDETWEKPLHHP 420
QY 367 EHLFDAQGHFVKPEAFPLPFSAGRRACLGEPFLARMELFLFTSLLOHFSFVSPTGQRRPSH 426
Db 421 GHFLDAEGRFVKREAFMPFSAGPRICLGEPLARMELFLFTSLLOHFSFVSPTGQRRPSD 480
QY 427 HGVFAFLYTPSPYELCAVPR 446
Db 481 RGAPYLWVLPSPYQLCAVLR 500

Search completed: February 25, 2004, 04:30:00
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 04:04:57 ; Search time 44 Seconds
(without alignments)
3198.207 Million cell updates/sec

Title: US-09-820-788A-2
Perfect score: 2330
Sequence: 1 MGLEALVPLAVIAIFLLLV.....HGVAFLVTPSPYELCAVPR 446

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL.25.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2283	98.0	500	4	Q16753
2	2197.5	94.3	497	6	Q8WNR5
3	2138.5	91.8	497	6	Q865W1
4	1769	75.9	500	6	Q29454
5	1751	75.2	500	6	Q9TJ4
6	1719.5	73.8	373	4	Q16804
7	1704.5	73.2	500	11	Q9JKY7
8	1698.5	72.9	500	11	Q91W87
9	1697	72.8	500	6	Q9TJ5
10	1633.5	70.1	500	11	Q8VCX0
11	1625	69.7	500	11	Q8CIM7
12	1622	69.6	500	11	Q9DBJ5
13	1600.5	68.7	504	11	Q921V1
14	1580.5	67.8	504	11	Q64529
15	1580.5	67.8	504	11	Q8BVD2
16	1567.5	67.3	504	11	Q64530

ALIGNMENTS

RESULT 1

Q16753 PRELIMINARY; PRT; 500 AA.
ID Q16753
AC Q16753;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Debrisoquine 4-hydroxylase mutant allele (CYP2D6-MAL).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Gonzalez F.J.;
RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; M33189; AAA35737.1; -;
DR HSSP; P00179; 1DT6.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 55899 MW; D5293E9BF74692C8 CRC64;

Query Match 98.0%; Score 2283; DB 4; Length 500;
Best Local Similarity 88.8%; Pred. No. 5.1e-179;
Matches 444; Conservative 1; Mismatches 1; Indels 54; Gaps 1;
Qy 1 MGLEALVPLAVIAIFLLLVLMHRRQWRWAARYSPGLPLPGIIGLLHVDFTONTPCFQD 60
DB 1 MGLEALVPLAVIAIFLLLVLMHRRQWRWAARYSPGLPLPGIIGLLHVDFTONTPCFQD 60


```
QY 1 MGLEALVPLAVTVAIFLLVLDLMHRRQRWAARYSPGPIPLPGIGNLLHVDFTONTYCDQ 60
Db 1 MGLDALVPLAVTVAIFLLVLDLMHRRQRWAARYSPGPIPLPGIGNLLHVDFTONTYCDQ 60
QY 61 LRRRGDVFSLQALMTVPVVLNGLAAVREALVTHGEDTADRPVPIQTILGFGPSSQ--- 117
Db 61 LRRRGDVFSLQALMTVPVVLNGLAAVREALVTHGEDTADRPVPIQTILGFGPSSQGVF 120
QY 118 -----GRPRPNGLDK 129
Db 121 LARYGPAWREORRFSVSTLNLGLCKSLQWVTEEAACLCFAFHSGRPRPNGLDK 180
QY 130 AVSNVIASITCGRRFEYDDPRFLRLDLAQEGKESGFLRVLNANVPVLLHHPALAGKV 189
Db 181 AVSNVIASITCGRRFEYDDPRFLRLDLAQEGKESGFLRVLNANVPVLLHHPALAGKV 240
QY 190 LRFQKAFUTQDDELTEHMTWDPAPPRDLTEAFLAEMKAKGNPESFNDENLRIVVA 249
Db 241 LRFQKAFUTQDDELTEHMTWDPAPPRDLTEAFLAEMKAKGNPESFNDENLRIVVA 300
QY 250 DLFSAGWTTSTTLAWGLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 309
Db 301 DLFSAGWTTSTTLAWGLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTTAVI 360
QY 310 HEVQRFQDVIPLGVTHMTSRDIEVQGFRIKPGTTLITNLSVVKDEAVWEKPFRRPHEHF 369
Db 361 HEVQRFQDVIPLGVTHMTSRDIEVQGFRIKPGTTLITNLSVVKDEAVWEKPFRRPHEHF 420
QY 370 LDAQHFVYKPEAFLEPSAGRRACLGELPLARMEFLFFTSLLQHFSSVPTGQPRSHGV 429
Db 421 LDAQHFVYKPEAFLEPSAGRRACLGELPLARMEFLFFTSLLQHFSSVPTGQPRSHGV 480
QY 430 FAFVTPSPSYELCAVPR 446
Db 481 FAFVTPSPSYELCAVPR 497

RESULT 4
Q29454
ID Q29454 PRELIMINARY; PRT; 500 AA.
AC Q29454;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome P-450 IID.
GN CYP2D.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]_TaxID=9913;
RP SEQUENCE FROM N.A.
RC STRAIN=PBVL 180; TISSUE=Liver;
RX MEDLINE=93011103; PubMed=1396678;
RA Tsuneoka Y., Matsuo Y., Higuchi R., Ichikawa Y.;
RT "Characterization of the cytochrome P-450IID subfamily in bovine
RT liver. Nucleotide sequences and microheterogeneity." ;
RL Eur. J. Biochem. 208:739-746 (1992).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; X68481; CAA48501.1; -.
DR PIR; S37284; S37284.
DR HSSP; P00179; 1DT6.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
```

```
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 55921 MW; 4F62F39050E2BED6 CRC64;

Query Match 75.9%; Score 1769; DB 6; Length 500;
Best Local Similarity 69.6%; Pred. No. 8.9e-137;
Matches 348; Conservative 39; Mismatches 59; Indels 54; Gaps 2;

QY 1 MGL---EALVPLAVTVAIFLLVLDLMHRRQRWAARYSPGPIPLPGIGNLLHVDFTONTYCV 57
Db 1 MGLSGDVLGLVALLFLLLDLMHRRSWAPYPPGPTPLPVLGNLLQVDFDPPRS 60
QY 58 FDLRRRGDVFSLQALMTVPVVLNGLAAVREALVTHGEDTADRPVPIQTILGFGPSSQ 117
Db 61 FDLRRRGDVFSLQALMTVPVVLNGLAAVREALVTHGEDTADRPVPIQTILGFGPSSQ 120
QY 118 -----GRPRPNGL 126
Db 121 GVILARYGDMAEQRRFSLTTLNFGKLSLEQWVTEEAACSCAFAQAGRPSPMDL 180
QY 127 LDKAVSNVIASITCGRRFEYDDPRFLRLDLAQEGKESGFLRVLNANVPVLLHHPALA 186
Db 181 LNKAVSNVIASITCGRRFEYDDPRFLRLDLAQEGKESGFLRVLNANVPVLLHHPALA 240
QY 187 GKVLRFQKAFUTQDDELTEHMTWDPAPPRDLTEAFLAEMKAKGNPESFNDENLR 246
Db 241 ARVPPAQKAFMALIDELAEQKMTDPTQPPRHLLTDAFLDEKVEKAGNPESFNDENLR 300
QY 247 VVADLFSAGWTTSTTLAWGLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTT 306
Db 301 VVADLFSAGWTTSTTLAWGLLMLHPDVQRRVQOEIDDDVIGVRRPEMGDQAHMPYTT 360
QY 307 AVIHEVQRFQDVIPLGVTHMTSRDIEVQGFRIKPGTTLITNLSVVKDEAVWEKPFRRP 366
Db 361 AVIHEVQRFQDVIPLGVTHMTSRDIEVQGFRIKPGTTLITNLSVVKDEAVWEKPFRRP 420
QY 367 EHFLDAQHFVYKPEAFLEPSAGRRACLGELPLARMEFLFFTSLLQHFSSVPTGQPRSH 426
Db 421 EHFLDAQHFVYKPEAFLEPSAGRRACLGELPLARMEFLFFTSLLQHFSSVPTGQPRSH 480
QY 427 HGVFAFLVTPSPSYELCAVPR 446
Db 481 HGVFAFLVTPSPSYELCAVPR 500

RESULT 5
Q29454
ID Q29454 PRELIMINARY; PRT; 500 AA.
AC Q29454;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome P450 2D/II.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]_TaxID=9986;
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand White; TISSUE=Liver;
RX MEDLINE=98391821; PubMed=972858;
RA Yamamoto Y., Ishizuka M., Takada A., Fujita S.;
RT "Cloning, tissue distribution, and functional expression of two novel
RT rabbit cytochrome P450 isozymes, CYP2D23 and CYP2D24." ;
RL J. Biochem. 124:503-508 (1998).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AB008785; BAA84473.1; -.
DR PIR; J0258; J0258.
DR HSSP; P00179; 1DT6.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
```

```
DR Pfam: PF00067; P450; 1.
DR PRINTS; PR01686; CYTOCHROME_P450.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 55604 MW; 434EC7C86BF6305B CRC64;

Query Match
Best Local Similarity 75.2%; Score 1751; DB 6; Length 500;
Matches 348; Conservative 31; Mismatches 67; Indels 54; Gaps 2;

QY 1 MGL---EALVPLAVIAVFLVLLVLMHRRQWAAAYSGPLPLPGLGNLLHVDFTQVPC 57
D 1 MGLSSEALAPLAVAVFLVLLVLMHRRQWAAAYSGPLPGLGNLLHVDFTQVPC 60
QY 58 FQLRRRFGDVSQQLAWTPVVVNLGLAAVREALVTHGSDTADRPVPIQTILGFGPRSQ 117
D 61 FQLRRRYGDVSQQLAWTPVVVNLGLAAVREALVTHGSDTADRPVPIQTILGFGPRSQ 120
QY 118 -----GRPRPGL 126
D 121 GVMARYGAWREQRFSVSTLRNFGLGKKSLEQWVTEATCLCAAFADHAGCPSPSML 180
QY 127 LDKAVSNVIASLTGRRFEDDPRFLRLDLAQEGKESGFLREVNLNAVVPVLLHHPALA 186
D 181 LNKAVCNVIASLTGCRFEDDPRFLRLDLAQEGKESGFLREVNLNAVVPVLLHHPALA 240
QY 187 GKVLRFQKFLQLDELTEHRTWDPAPPRDLTEAFLAEMEAKGNPESFNENLRI 246
D 241 DKVFGQKAFMALDELTEHRTWDPAPPRDLTEAFLAEMEAKGNPESFNENLRI 300
QY 247 VVADLFSAGMWTSTTLANGLLMLHPDVQRRVQOEIDDDVIGQVRRPEMGDQAMPTT 306
D 301 VVTDLFAAGMWTSTTLANGLLMLHPDVQRRVQOEIDDDVIGQVRRPEMGDQAMPTT 360
QY 307 AVIHEVQRRGDIPLVGLVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPRFHP 366
D 361 AVIHEVQRRGDIPLVGLVTHMTSRDIEVQGFRIPKGTTLITNLSSVLKDEAVWEKPRFHP 420
QY 367 EHFLDAQGHFVKPEAFPLPSAGRAACLGELPLARMELFFFTSLLOHFSFVPTGQPRPSH 426
D 421 GHFLDAQGHFVKPEAFPLPSAGRAACLGELPLARMELFFFTSLLOHFSFVPTGQPRPSH 480
QY 427 HGVFAFLVTPSPYELCAVPR 446
D 481 QGAPATLVTPSPYELCAVPR 500

RESULT 6
ID Q16804 PRELIMINARY; PRT; 373 AA.
AC Q16804;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE Cytochrome P450db1 (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89155788; PubMed=2466049;
RA Manns M.P., Johnson E.F., Griffin K.J., Tan E.M., Sullivan K.F.;
RT "Major antigen of liver kidney microsomal autoantibodies in idiopathic
RT autoimmune hepatitis is cytochrome P450db1."
RL J. Clin. Invest. 83:1066-1072 (1989).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; M24499; AA336403.1; -.
DR HSSP; P00179; 1DT6.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.

DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PROSITE; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
KW Heme; Monooxygenase; Oxidoreductase.
FT NON TER 1
SQ SEQUENCE 373 AA; 42005 MW; 9FDEDG7B0BA487A4 CRC64;

Query Match
Best Local Similarity 73.8%; Score 1719.5; DB 4; Length 373;
Matches 333; Conservative 1; Mismatches 3; Indels 19; Gaps 1;

QY 110 LGFGPRS-----QGRPRPGLLDKAVSNVIASLTGRRFEDDPR 150
D 18 LGLGKKSLEQWVTEEAACLCFAFNHSGRPRPGLLDKAVSNVIASLTGRRFEDDPR 77
QY 151 FLRLDLAQEGKESGFLREVNLNAVVPVLLHHPALAGKVLRFQKFLQLDELTEHRT 210
D 78 FLRLDLAQEGKESGFLREVNLNAVVPVLLHHPALAGKVLRFQKFLQLDELTEHRT 137
QY 211 WDPAPPRDLTEAFLAEMEAKGNPESFNENLRIVVADLFSAGMWTSTTLANGLLML 270
D 138 WDPAPPRDLTEAFLAEMEAKGNPESFNENLRIVVADLFSAGMWTSTTLANGLLML 197
QY 271 ILHPDVQRRVQOEIDDDVIGQVRRPEMGDQAMPTTAVIHEVQRRGDIPLVGLVTHMTSRD 330
D 198 ILHPDVQRRVQOEIDDDVIGQVRRPEMGDQAMPTTAVIHEVQRRGDIPLVGLVTHMTSRD 257
QY 331 IEVQGFRIPKGTTLITNLSSVLKDEAVWEKPRFHPPEHFLDAQGHFVKPEAFPLPSAGRR 390
D 258 IEVQGFRIPKGTTLITNLSSVLKDEAVWEKPRFHPPEHFLDAQGHFVKPEAFPLPSAGRR 317
QY 391 ACLGEPLARMELFFFTSLLOHFSFVPTGQPRPSHGVFAFLVTPSPYELCAVPR 446
D 318 ACLGEPLARMELFFFTSLLOHFSFVPTGQPRPSHGVFAFLVTPSPYELCAVPR 373

RESULT 7
ID Q9JKY7 PRELIMINARY; PRT; 500 AA.
AC Q9JKY7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE Cytochrome P450 CYP2D22.
GN CYP2D22.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=20485155; PubMed=11032406;
RA Blume N., Leonard J., Xu Z.J., Watanabe O., Remotti H., Fishman J.;
RT "Characterization of Cyp2d22, a novel cytochrome P450 expressed in
RT mouse mammary cells."
RL Arch. Biochem. Biophys. 381:191-204 (2000).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF221525; AAF34652.1; -.
DR HSSP; P00179; 1DT6.
DR MGD; MGI:1929474; Cyp2d22.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 56467 MW; FACB35854CBC3F1F CRC64;
```



```
DR Pfam: PF00067; p450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
SQ SEQUENCE 500 AA; 56975 MW; B3DDCD88DFA3F265 CRC64;

Query Match 69.7%; Score 1625; DB 11; Length 500;
Best Local Similarity 63.4%; Pred. No. 6e-125;
Matches 317; Conservative 49; Mismatches 80; Indels 54; Gaps 2;

QY 1 MGL---EALVPLAVIAIFLLVLDLHRRQRWAARYSGPLPLPGLGNLLHVDVQNTPYC 57
DB 1 MGLVGGDLDWAVVITFAIFLLVLDLHRRQRWTACYPGVPFPPGLGNLLQVDFENIPYS 60

QY 58 FDLRRRFGDVSFSLQALWTPVVVINGLAARVLAETHGSDTADRPVPTQLIGFGRSQ 117
DB 61 FYKLQNRVGNVFSQMAWKPVVVVNGLVKAVRELLVYGEDTSRPLMPYINHYGYSK 120

QY 118 G-----RPFPPNGL 126
DB 121 GVILAPYGPWEQRFRFSVSTLRDFGLGKKSLEQWVTEAGHLCDAFTEAEHPNPSPL 180

QY 127 LDKAVSNVLTASLCGRFEYDDPRFLRLDLAQEGLKEESGFLREVLNAVPLVLLHPALA 186
DB 181 LSKAVSNVLTASLIYARRFEYEDPFNRMLTKLESIGEDTGFVGEVLNAIPMLLHPGLP 240

QY 187 GKVLRFQKAFLTQDLDELTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPSESSFNENLRI 246
DB 241 DKAFPKLNSFIALVKNMLTEHSDTWDPAQPPRDLTDAFLAEVEKAKGNPSESSFNENLRI 300

QY 247 VVADLFSAGMVTSTTTLAWGLLMLHPDVQRRVQOEIDVIGQVRRPEMGDAHMPYTT 306
DB 301 VVIDLFMAGMVTSTTSLWALLMLHPDVQRRVQOEIDVIGHVHRPEMAQARMPYTN 360

QY 307 AVIHEVQRGDIPLGVTHMTSRDIEVQGRIPKGTTLITNLSVLKDAVMEKPPRFP 366
DB 361 AVIHEVQRADIVPTNLPMTSRDIKFQDFIPKGTTLIPNLSVLKDETVWEKPLRFP 420

QY 367 EHLDAQGHFVKPEAFPSAGRRACLGELPLARMELFFFTLLQHFSSVPTGQPRPSH 426
DB 421 EHLDAQGHFVKHEAFMPFSAGRRSCLGELPLARMELFFFTLLQHFSSVPTGQPRPSD 480

QY 427 HGVEAFVLTPTSPVELCAVR 446
DB 481 YGIYTPVTPPEYQLCAVAR 500

RESULT 13
Q921V1 PRELIMINARY; PRT; 504 AA.
AC Q921V1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
```

```
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustingich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AK004915; BAB23666.1; -.
DR HSSP; P00179; 1D76.
DR MGD; MGI:1923529; Cyp2d26.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 56949 MW; F4C9A03E04C8752D CRC64;

Query Match 69.6%; Score 1622; DB 11; Length 500;
Best Local Similarity 63.4%; Pred. No. 1.1e-124;
Matches 317; Conservative 48; Mismatches 81; Indels 54; Gaps 2;

QY 1 MGL---EALVPLAVIAIFLLVLDLHRRQRWAARYSGPLPLPGLGNLLHVDVQNTPYC 57
DB 1 MGLVGGDLDWAVVITFAIFLLVLDLHRRQRWTACYPGVPFPPGLGNLLQVDFENIPYS 60

QY 58 FDLRRRFGDVSFSLQALWTPVVVINGLAARVLAETHGSDTADRPVPTQLIGFGRSQ 117
DB 61 FYKLQNRVGNVFSQMAWKPVVVVNGLVKAVRELLVYGEDTSRPLMPYINHYGYSK 120

QY 118 G-----RPFPPNGL 126
DB 121 GVILAPYGPWEQRFRFSVSTLRDFGLGKKSLEQWVTEAGHLCDAFTEAEHPNPSPL 180

QY 127 LDKAVSNVLTASLCGRFEYDDPRFLRLDLAQEGLKEESGFLREVLNAVPLVLLHPALA 186
DB 181 LSKAVSNVLTASLIYARRFEYEDPFNRMLTKLESIGEDTGFVGEVLNAIPMLLHPGLP 240

QY 187 GKVLRFQKAFLTQDLDELTEHRMTWDPAPPRDLTEAFLAEMEKAKGNPSESSFNENLRI 246
DB 241 DKAFPKLNSFIALVKNMLTEHSDTWDPAQPPRDLTDAFLAEVEKAKGNPSESSFNENLRI 300

QY 247 VVADLFSAGMVTSTTTLAWGLLMLHPDVQRRVQOEIDVIGQVRRPEMGDAHMPYTT 306
DB 301 VVIDLFMAGMVTSTTSLWALLMLHPDVQRRVQOEIDVIGHVHRPEMAQARMPYTN 360

QY 307 AVIHEVQRGDIPLGVTHMTSRDIEVQGRIPKGTTLITNLSVLKDAVMEKPPRFP 366
DB 361 AVIHEVQRADIVPTNLPMTSRDIKFQDFIPKGTTLIPNLSVLKDETVWEKPLRFP 420

QY 367 EHLDAQGHFVKPEAFPSAGRRACLGELPLARMELFFFTLLQHFSSVPTGQPRPSH 426
DB 421 EHLDAQGHFVKHEAFMPFSAGRRSCLGELPLARMELFFFTLLQHFSSVPTGQPRPSD 480

QY 427 HGVEAFVLTPTSPVELCAVR 446
DB 481 YGIYTPVTPPEYQLCAVAR 500

RESULT 13
Q921V1 PRELIMINARY; PRT; 504 AA.
AC Q921V1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
```

```
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DE 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Cytochrome P450, 2d9.
GN CYP2D9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; BC010593; AAH10593.1; -.
DR MGD; MGI:88606; Cyp2d9.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; P:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 504 AA; 56950 MW; 6DC93B3985EFB8A2 CRC64;

Query Match 68.7%; Score 1600.5; DB 11; Length 504;
Best Local Similarity 62.6%; Pred. No. 6.3e-123;
Matches 308; Conservative 53; Mismatches 80; Indels 51; Gaps 1;

QY 6 LVPLAVIAIFLLVLDLHRRORWAARYSPGLPLPGIIGNLLHVDFTQNTYCFDOLRRRF 65
DB 9 LMPVAIFTVIFLLVLDLTHQRWTSRYPGVPWPVIGNLLQVDLGNMPSYLYKQRY 68
QY 66 GDVFSQLAWTPVVVINGLAAREALVTHGSDTADRPVPVITQILGFGPRSQ----- 117
DB 69 GDVFSQAWKPIVINGLAKEMKMLTCCGDTADRPVPVIFEYLGKPGSGGVVLPYG 128
QY 118 -----GRPPFNGLLDKAVSNV 134
DB 129 PEWREQRFRFSVTLRNFGLGKKSLEDVWTKEARHLCDAFQAAGRSINPNTMLNKSTCNV 188
QY 135 IASITCGRRFEYDDPRFLRLDLAQEGLEKESGFLREVINAAPVLLHHPALAGKVLRFQK 194
DB 189 IASLIFARRFEYEDPFLIRMLKVLQESLVEISGLIPEVINAPFILLRIPRLADKALQOK 248
QY 195 AFLTQDLDELTEHRMTWDPAPPRDLTEAFLAEKAKGNPSSFNENLRIVVADLFSA 254
DB 249 SFNAILDNLNLTNTTWDVPVQAPRNLTDAFLAEIKAKGNPSSFNENLRIVVADLFSA 308
QY 255 GMVTTSTTLAAGLLMLHPDVQRVQOEIDVIGQVRRPEMGDQAHMPYTTAVIHEVOR 314
DB 309 GMLATSTTSLWALMLHPDVQRVQOEIDVIGQVRRPEMGDQAHMPYTTAVIHEVOR 368
QY 315 FGDIVPLGVTHMTSDIEVOGFRIKPGTTLITNLSVLKDEAVWEKPRFHPHFDAQ 374
DB 369 FGDIVPNLPRITSHDIEVDQFLIPKGTILLPNMSSMLKDSVWEKPLRFHPHFDAQ 428
QY 375 HFVKEAFPLPFSAGRRACLGELPRLMELFFFTSLQHFSSVPTGQPRPSHHGVFAFLV 434
DB 429 HFVKEAFMPFSAGRRSCLGELARMLFLFTCLLQRFSSVDPGQPRPSHHGVFAFLV 488
QY 435 TPSPYELCAVR 446
DB 489 APSPYQLCAVR 500

RESULT 14
Q64529 PRELIMINARY; PRT; 504 AA.
AC Q64529;
DT 01-NOV-1996 (TREMELrel. 01, Created)
```

Search completed: February 25, 2004, 04:30:59
Job time : 48 secs

```
RESULT 15
Q8BVD2      PRELIMINARY;      PRT;      504 AA.
ID          Q8BVD2;
AC          Q8BVD2;
DT          01-MAR-2003 (TEMBLrel. 23, Created)
DT          01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT          01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE          Cytochrome P450.
OS          Mus musculus (Mouse).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX          NCBI_TaxID=10090;
RN          [1]
SEQUENCE FROM N.A.
RP          STRAIN=C57BL/6J; TISSUE=Colon;
RC          MEDLINE=22354683; PubMed=12466851;
RA          The FANTOM Consortium.
RA          the RIKEN Genome Exploration Research Group Phase I & II Team;
RT          "Analysis of the mouse transcriptome based on functional annotation of
RL          60,770 full-length cDNAs.";
RL          Nature 420:563-573 (2002).
DR          EMBL; AK078880; BAC37440.1; -.
DR          GO: 0016712; F:oxidoreductase activity, acting on paired d. . . ; IEA.
DR          GO: 0006118; P:electron transport; IEA.
DR          InterPro; IPR001128; Cytochrome P450.
DR          InterPro; IPR008069; EP450_CYP2D.
DR          Pfam; PF00067; P450; 1.
DR          PRINTS; PR01686; EP450ICYP2D.
DR          PRINTS; PR00385; P450.
DR          PROSITE; PS00086; CYTOCHROME_P450; 1.
SQ          SEQUENCE 504 AA; 57743 MW; 4B29AB3160997957 CRC64;

Query Match      67.8%; Score 1580.5; DB 11; Length 504;
Best Local Similarity 62.6%; Pred. No. 2.8e-121;
Matches 306; Conservative 51; Mismatches 81; Indels 51; Gaps 1;

QY 9 LAVIVAIFLLVDLMHRRGRMAARYSPGELPLPGLGNLLHVFQNTPTPCFDQLRRRFGDV 68
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
12 VAIFTVIFILLVDLMHRRQSWTSCYPPGPVPVPLGNLLQVDLNNMPSLYKLQNRYGDV 71
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
69 FSIQLAWTPVVLNGLAAREALVTHGEDTADRPVPIITQILGFGPRSGRPP----- 121
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
72 FSIQMAKMPVWVNRMKAMKVELLTCGEDTADRPVPIFEHLGFKPRSQGMIFAPYGPW 131
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 -----RPNGLDKAVSNVIAS 137
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
132 REQRRFSLSLRNFGLRKSLSEWVIKEAGHLCDFTTQAGQYINENTMLKATCNVIAS 191
QY 138 LTCGRFXYDDPRFLDLDAQGLKESGFLREVLNAVVPVLHHPALAGKVLRFQKAF 197
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
192 LIPARRFEYEDPYLIRMLKVLDSLTSLGLIPEVINTFPILLHIPLADKFLQSKSFI 251
QY 198 TQDELLETHRTWTDPAQPRDLTEAFLEMEKAKGNPSSFNENLRIWVADLPSAGMV 257
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
252 AIVDNLLTENRTTWDPAQSPRNLTDAFLAEIEKAKGNPSSFNENLRMVVIDLFTAGIL 311
QY 258 TTSITLAWGLLMLHPDVQRVQQEIDDVIGVRRPEMGDQAHMPYTTAVIHEVORFGD 317
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
312 TTSITLAWGLLMLHPDVQRVQQEIDDVIGVRRPEMGDQAHMPYTTAVIHEVORFGD 371
QY 318 IVPGLVTHMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPPRFPHEFLDAQGHFV 377
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
372 IVPGLVTHMTSRDIEVQGFRIKPGTTLITNLSSVLKDEAVWEKPPRFPHEFLDAQGHFV 431
QY 378 KPEAFILPFSAGRACLGELARMEFLFTTSLLOHFSFVPTGQPRPSHGVFAFLVTPS 437
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
432 KHEAFITFSAGRACLGELARMEFLFTTCLLQRFSSVSPDQGPQPSDRHVFVSIWVAPS 491
QY 438 PYELCAVPR 446
Db : : : : :
492 PYOLCAVIR 500
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 29, 2004, 00:03:55 ; Search time 2375 Seconds
(without alignments)
5607.803 Million cell updates/sec

Title: US-09-820-788A-2
Perfect score: 2330
Sequence: 1 MGLEALVPLAVIAVIFLLLV.....HGVFAFLVTPSPYELCAVPR 446

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DRV=xlh
-Q=/cgn2_1/USPTO_spool/US09820788/runat_24022004.141408_9107/app.query.fasta_1.583
-DB=EST -Qfmt=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09820788 @CNC 1.1 2135 @runat_24022004.141408_9107 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST :
1: em_estba :
2: em_esthum :
3: em_estin :
4: em_estmu :
5: em_estov :
6: em_estpl :
7: em_estro :
8: em_htc :
9: gb_est1 :
10: gb_est2 :
11: gb_htc :
12: gb_est3 :
13: gb_est4 :
14: gb_est5 :
15: em_estfun :
16: em_estom :
17: em_gss_hum :
18: em_gss_inv :
19: em_gss_pln :
20: em_gss_vrt :
21: em_gss_fun :
22: em_gss_mam :
23: em_gss_mus :
24: em_gss_pro :
25: em_gss_rod :
26: em_gss_pbg :
27: em_gss_vrl :
28: gb_gss1 :

29: gb_gss2 :*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1622	69.6	1646	11	AK004915	AK004915 Mus muscu
2	1609	69.1	1201	13	BX422591	BX422591 BX422591
3	1603	68.8	2429	11	AK090296	AK090296 Mus muscu
4	1580.5	67.8	1624	11	AK078880	AK078880 Mus muscu
5	1304	56.0	1497	11	BC051039	BC051039 Mus muscu
6	1246.5	53.5	1201	13	BX422592	BX422592 BX422592
7	1228.5	52.7	814	14	CK032845	CK032845 AGENCOURT
8	1228	52.7	798	14	CF271947	CF271947 AGENCOURT
9	1227.5	52.7	803	14	CK032848	CK032848 AGENCOURT
10	1225	52.6	3079	11	AK004933	AK004933 Mus muscu
11	1223	52.5	767	14	CK032844	CK032844 AGENCOURT
12	1221.5	52.4	780	14	CK032843	CK032843 AGENCOURT
13	1214	52.1	804	14	CK032846	CK032846 AGENCOURT
14	1207.5	51.8	1088	13	BX432878	BX432878 BX432878
15	1201.5	51.6	982	12	BG743095	BG743095 602634230
16	1196	51.3	807	14	CK032847	CK032847 AGENCOURT
17	1188.5	51.0	796	14	CK130223	CK130223 AGENCOURT
18	1179.5	50.6	791	14	CF271949	CF271949 AGENCOURT
19	1143.5	49.1	1071	13	BX432877	BX432877 BX432877
20	1116.5	47.9	902	13	CK030832	CK030832 BX430832
21	1095.5	47.0	774	14	CK032849	CK032849 AGENCOURT
22	1065	45.7	1805	11	AK004984	AK004984 Mus muscu
23	1047.5	45.0	962	13	CK030831	CK030831 BX430831
24	1014	43.5	880	12	BI217581	BI217581 602934015
25	1011	43.4	586	14	CB162434	CB162434 K-EST0223
26	1004	43.1	631	14	CB456776	CB456776 714071 MA
27	993	42.6	808	12	BI328628	BI328628 602984542
28	992	42.6	798	12	BI247704	BI247704 602959323
29	991	42.5	796	9	AI255822	AI255822 ui188f05.x
30	983	42.2	809	9	AI987772	AI987772 um06906.x
31	980	42.1	788	9	AI266991	AI266991 u109604.x
32	975.5	41.9	969	12	BI217057	BI217057 602933363
33	974	41.8	1066	14	CB195223	CB195223 AGENCOURT
34	961.5	41.3	787	14	CK130220	CK130220 AGENCOURT
35	961	41.2	802	12	BG775066	BG775066 602650030
36	950	40.8	725	10	BF533324	BF533324 602073784
37	948.5	40.7	820	12	BI331562	BI331562 602960980
38	947.5	40.7	793	9	AI256232	AI256232 ui196f03.x
39	946	40.6	736	12	BI103579	BI103579 602888726
40	944	40.5	798	9	AI194824	AI194824 ui157c03.x
41	942.5	40.5	816	14	CK130224	CK130224 AGENCOURT
42	941.5	40.4	754	14	CK130221	CK130221 AGENCOURT
43	936	40.2	782	9	AI326364	AI326364 mm10c06.x
44	935	40.1	630	14	CF360997	CF360997 827299 MA
45	931	40.0	803	14	CK130222	CK130222 AGENCOURT

ALIGNMENTS

RESULT 1
AK004915
LOCUS
DEFINITION
Mus musculus adult male liver cDNA, RIKEN full-length enriched library, clone:1300006E06 product:CYTOCHROME P450 2D2 (EC 1.14.14.1) (CYP11D2) (P450-DB2) (P450-CMF2) (DEBRISOQUINE 4-HYDROXYLASE) homolog [Rattus norvegicus], full insert sequence.
ACCESSION AK004915
VERSION AK004915.1 GI:12836460
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; 1646 bp mRNA linear HTC 20-SRP-2003


```
QY 118 ----- 118
Db 494 ACCCTCGTGAATTCGGCTTGGGCAAAATCACTGGAGCAGTGGGTGACAGAGGAGCT 553
QY 119 -----
Db 554 GGCACCTCTGGCATGCTTTCCACCAAGGAGCGGCAACATCCCTTCAATCCGAGCCCTC 613
QY 127 LeuAspLysAlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyr 146
Db 614 CTAAGTAAAGCTGTGAGCAAGTGATCGCTCCCTCATTTATGCTCGTGAATTTGATAT 673
QY 147 AspAspProArgPheLeuArgLeuLeuAspLeuAlaGlnGluClyLeuLysGluGluSer 166
Db 674 GAAGACCCCTTCTTCAACAGGATGCTCAAAACATTGAAGAAAGCTTAGGAGAAGCACT 733
QY 167 GlyPheLeuArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAla 186
Db 734 GGCTTCTGGAGAGTGCTGAATGCTCCCAATGCTACTACACATCCCTGGTTGGCT 793
QY 187 GlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGlu 206
Db 794 GATAAAGCTTCCCAAGCTGAATTCATTATAGCTTAGTGAATGAAGATGTTAATTGAG 853
QY 207 HisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAla 226
Db 854 CAGACTCGACCTGGGATCCTGCACAGCCACCCGAGATTTGACTGATGCTTCTGGCT 913
QY 227 GluMetGluLysAlaLysGlyAsnProGluSerSerPheAsnAspGluLeuArgIle 246
Db 914 GAGTGGAGAGGCCAAGGGGAATCTTGAGAGCAGCTTCAATGATGAAGAACCTTGGGCATA 973
QY 247 ValValAlaAspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTrpGly 266
Db 974 GTGTGATTGACCTGTTTCATGCGAGGATGTTGACCACTCAACACACTGCTCTGGCC 1033
QY 267 LeuLeuLeuMetIleLeuHisProAspValGlnArgValGlnGlnGluLeuAsp 286
Db 1034 CTGCTGCTCATGATCTGCATCCAGATGTGCAGCGCGTGTCCACGAGAAATCGACGAG 1093
QY 287 ValIleGlyClnValArgProGluMetGlyAspGlnAlaHisMetProTyrThrThr 306
Db 1094 GTCATAGGCGACGTGCGGATCCAGAGATGGCAGACCGCGCATGCCCTTACACAAAC 1153
QY 307 AlaValIleHisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMet 326
Db 1154 GCGGTCAATTCATGAGTGCACGCTTTCAGACATCTCTCCCAACAAATTTACACATATG 1213
QY 327 ThrSerArgAspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThr 346
Db 1214 ACATCCGAGACATTAAATTCGAAGACTTCTTCATCCCAAGGGGACGACACTCATTTCC 1273
QY 347 AsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisPro 366
Db 1274 AACCTGTCTCGTGTGAAGATGATGATGTGTGGGAAGACCCCTCGCTTCTATCTC 1333
QY 367 GluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSer 386
Db 1334 GAACACTTCTGGATGCCAGGGCCACTTGTGTGAGCAGCAGGCTTTCATGCACTTCTCA 1393
QY 387 AlaGlyArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhe 406
Db 1394 GCAGGCGCAGATCATGCTCGGGGAGCCCTTGGCCGCGATGGAGCTTCTCTTCTTTC 1453
QY 407 ThrSerLeuLeuGlnHisPheSerPheSerValProThrGlyClnProArgProSerHis 426
Db 1454 ACTGTCTCTCGAGCGCTTATGCTTCTCAGTGCCCGATGGACAGCCAGCCAGCTGAT 1513
QY 427 HisGlyValPheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1514 TATGGCATCTATACAATGCCAGTTACTCCAGAGCCCTATCAGCTCTGTGAGTGGCTCGA 1573
```

RESULT 2

BX422591/c
LOCUS
DEFINITION
BX422591 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CS0DM008Y007 3-PRIME, mRNA sequence.
ACCESSION
BX422591
VERSION
BX422591.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10301.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DM008AH04NP1&cluster=10301.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DM008AH04NP1.

FEATURES
source
1..1201

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM008Y007"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN

Alignment Scores:

Pred. No.: 1,76e-164 Length: 1201
Score: 1609.00 Matches: 318
Percent Similarity: 96.66% Conservative: 0
Best Local Similarity: 96.66% Mismatches: 9
Query Match: 69.06% Indels: 3
DB: 13 Gaps: 0

US-09-820-788A-2 (1-446) x BX422591 (1-1201)

QY 120 PropheArgProAsnGlyLeuLeuAspLysAlaValSerAsnValIleAlaSerLeuThr 139
Db 998 CCGTTCGCCCCCAACGGTCTTTK-GAMAAAACCGTGAGCAACGTCATSTCTTCCCTCMAC 940
QY 140 CysGlyArgArgPheGluTyrAspAspProArgPheLeuArgLeuLeuAspLeuAlaGln 159
Db 939 TCGCGGCGCCGCTTCGAGTACGACACCCCTCGCTTCCTCAGGCTGCTGGACCTAGCTCAG 880
QY 160 GluGlyLeuLysGluGluSerGlyPheLeuArgGluValLeuAsnAlaValProValLeu 179
Db 879 GAGGAGCTGAGGAGAGTTCGGGCTTTCTGCGGAGGTGCTGAATGCTGTCCTCCCTC 820
QY 180 LeuHisIle-ProAlaLeuAlaGlyLysValLeuArgPhe-GlnLysAlaPheLeuThrG 199
Db 819 CTGCATATCCCGAGCGCTGGCTGGCAAGTCTCTACGCTCCMAAAGGCTTCTTCTGACCC 760
QY 199 InLeuAspGluLeuLeuThrGluHisArgMetThrTrpAspProAlaGlnProProArg 219
Db 759 AGCTGGATGAGTGTAACTGAGCACAGATGACCTGGAGACCCAGCCGCCGCCCGCAG 700

```
Qy 219 spLeuThrGluAlaPheLeuAlaGluMetGluLeuAlaLysGlyAsnProGluSerSerP 239
Db 699 ACTGACTGAGCCCTTCTTGGCAGAGATGAGAGCCCAAGGGGAACCTTGAGACAGCT 640
Qy 239 heAsnAspGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetValThrT 259
Db 639 TCAATGATGAGAACCTTGGCAGATAGTGGTGGCTGACCTGTTCTTCCGGGATGGTGACCA 580
Qy 259 hrSerThrThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGlnArgA 279
Db 579 CCTGACCAACGCTGGCTGGGCGCTCTCTCATGATCCGATCCGATGTGCAGCGCC 520
Qy 279 rGValGlnGlnLeuLeuAspValIleGlyGlnValArgArgProGluMetGlyAspG 299
Db 519 GTGTCCACAGAGATCGACGACGTAGTAGGCGAGTGGCGGACGAGATGGGTGACC 460
Qy 299 lnAlaHisMetProTyrThrThrAlaValIleHisGluValGlnArgPheGlyAspIleV 319
Db 459 AGGCTCACATGCCCTACACCACTCCGCTGATTCATGAGTGCAGCGCTTTGGGGACATCG 400
Qy 319 alProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPheArgIleP 339
Db 399 TCCCCCTGGGTGACCCATATGACATCCCGTGACATCGAAGTACAGGGCTTCGCGATCC 340
Qy 339 rGlyGlyThrThrLeuIleThrAsnLeuSerValLeuLysAspGluAlaValTrpG 359
Db 339 CTAAAGGAACGACATCATCACCACTGTCTGCTGCTGAGGATGAGCGCGTCTGG 280
Qy 359 LuLeuProPheArgPheHisProGluHisPheLeuAspAlaGlnGlyHisPheValLysP 379
Db 279 AGAAGCCCTTCGCTCCACCCGGAACACTTCTTGATGCCAGGGCCACTTTGTGAAGC 220
Qy 379 rGluAlaPheLeuProPheSerAlaGlyArgArgAlaCysLeuGlyGluProLeuAlaA 399
Db 219 CGAGGCGCTTCTGCTTCTCAGCAGCGCGCGTGCATGCTCGGGGAGCGCCCTGGGCC 160
Qy 399 rGMetGluLeuPheLeuPheThrSerLeuLeuGlnHisPheSerPheSerValProT 419
Db 159 GCATGGAGCTTCTCTTCTTCACTCTCTGCTGAGCAGCTTCAGCTTCTCGGTGCCCA 100
Qy 419 hrGlyGlnProArgProSerHisHisGlyValPheAlaPheLeuValThrProSerProT 439
Db 99 CTGGACAGCGCGCGCCGACCACTGTTCTTCTGCTTCTGCTGAGCCCATCCCCCT 40
Qy 439 YrGluLeuCysAlaValProArg 446
Db 39 ATGAGCTTTGTGCTGTGCCCCGC 17

RESULT 3
ACCESSION AK090296
VERSION AK090296.1 GI:26105861
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
JOURNAL Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
MEDLINE Meth. Enzymol. 303, 19-44 (1999)
PUBMED 99279253
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
```

```
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Izawa, Y., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Yoneda, Y., Ishikawa, T., Ogawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexed sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
PUBMED 11076861
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12401400
REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imorani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koyama, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki, A.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
FEATURES
source Location/Qualifiers
1..2429
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="PANTOM_DB:G630039L19"
/db_xref="MGI:2429202"
/db_xref="taxon:10090"
/clone="G630039L19"
/tissue type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="21 days neonate"
misc_feature 1..1410
/note="cytochrome P450, 2d22 (MGI:MGI:1929474,
GB|NM_019823, evidence: BLASTN, 99%, match=1477)
putative"
polyA_signal 2404..2409
/note="putative"
```

polyA_site
2429
/note="putative"

ORIGIN

Alignment Scores:
Pred. No.: 2,53e-163 Length: 2429
Score: 1603.00 Matches: 318
Percent Similarity: 73.99% Conservative: 29
Best Local Similarity: 67.80% Mismatches: 71
Query Match: 68.80% Indels: 52
DB: 11 Gaps: 1

US-09-820-788A-2 (1-446) x AK090296 (1-2429)

Qy 29 TTPAlaAlaArgTyrSerProGlyProLeuProGlyLeuGlyAsnLeuHis 48
|||
Db 3 TGGACTGCCACTACCTCCAGCCCTATGCGTGGCTGTCTGGGAACCTGCTGCAC 62
|||
Qy 49 ValAspPheGlnAsnThrProTyrCysPheAspGlnLeuArgArgPheGlyAspVal 68
:::|||||
Db 63 ATGAGCTTCAGAAATATGCGAGCGCTTCCAAAGAGCTCGCGGGTCGTATGGGGAGCTG 122
|||
Qy 69 PheSerLeuGlnLeuAlaTrpThrProValValValLeuAsnGlyLeuAlaAlaValArg 88
:::|||||
Db 123 TTGAGCTACAGCTGCGCTCTGAGTCAGTGGTGTACTAAATGGCTGACGCGCTCGGA 182
|||
Qy 89 GluAlaLeuValThrHisGlyGluAspThrAlaAspArgProProValProIleThrGln 108
|||
Db 183 GAGGCACTGGTGAACACACAGCAGGACACTGCTGACCGGCCACCGCTCAITTCATGAC 242
|||
Qy 109 IleLeuGlyPheGlyProArgSerGln 117
:::|||||
Db 243 CTGCTGGGCTTTGGACACCGCTCTCAAGGTATAGTCTTAGCAGCGTATGGGCTGCGCTGG 302
|||
Qy 117 ----- 117
Db 303 CGTCAGAGCGGCGCTTCTGTGTCTACCATGACCACTTTGGCGCTGGGCAAGATCA 362
|||
Qy 117 ----- 117
Db 363 CTGAGCAGTGGGTGACTGAGGAGCCAGATCCCTCTGTGCGCTTCGCTGACCATACT 422
|||
Qy 118 GlyArgProPheArgProGlnGlyLeuLeuAspLysAlaValSerAsnValIleAlaSer 137
|||
Db 423 GGACACCTTTACGCCCTAACACCTTATGGACAAAGCAGTGTGTAACTGATCGCGTCC 482
|||
Qy 138 LeuThrCysGlyArgArgPheGluTyrAspAspProArgPheLeuArgLeuAspLeu 157
|||
Db 483 CTCTCTATGCTGCGCTTTGAGTACGATGACCCACGCTTCATCAGGCTACTGGGCTTG 542
|||
Qy 158 AlaGlnGluGlyLeuLysGluLysSerGlyPheLeuArgGluValLeuAsnAlaValPro 177
:::|||||
Db 543 TTTAGAGAACTTTTAAAGAGAGAGCTGATTCCTACCCATGTTCTGAAATGTTCCCG 602
|||
Qy 178 ValLeuLeuHisIleProAlaLeuAlaGlyLysValLeuArgPheGlnLysAlaPheLeu 197
:::|||||
Db 603 ATGCTCTGCGCATCCCGGGCTGTTGGCAAGGCTCTTCCCTGGGAAAGGCGCTTTGTT 662
|||
Qy 198 ThrGlnLeuAspGluLeuLeuThrGluHisArgMetThrTrpAspProAlaGlnProPro 217
:::|||||
Db 663 ACCATGTTGATGAGCTGCTGGCTGAACACAGACGACCTCGGGACCTTACCGACCGCCCG 722
|||
Qy 218 ArgAspLeuThrGluAlaPheLeuAlaGluMetGluLysAlaLysGlyAsnProGluSer 237
:::|||||
Db 723 A-GATTGTAATGATGCTTCTGCTGCTGAGTGGAGAGGCCACAGGGGAATCTTGAGAGC 780
|||
Qy 238 SerPheAsnAspGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetVal 257
|||
Db 781 AGCTTCAATGATGAGAACCTGCGACGGTAGTGGGTGACCTGTTCTCTCGAGGAGTGGTG 840
|||
Qy 258 ThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGln 277
:::|||||
Db 841 ACCACTTCAACCACTGCTGGGCGCTGATGCTCATGATCTCTCATCCAGATGTGCGAG 900
|||

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861 4
TITLE JOURNAL REFERENCE AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 695-690 (2001) 5
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL REFERENCE AUTHORS	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 1624)
TITLE JOURNAL REFERENCE AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanaoka, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE JOURNAL	Direct Submission Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers
FEATURES source	1..1624 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM DB:9030605E09" /db_xref="MGI:2392571" /db_xref="taxon:10090" /clone="9030605E09" /sex="male" /tissue type="colon" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 50..1564 /note="unnamed protein product; CYTOCHROME P450, 2D9 homolog [Mus musculus] (SPTR Q921V1, evidence: FASTY, 87.7%ID, 100%length, match=1512) putative" /codon_start=1 /protein_id="BAC37440.1" /db_xref="GI:26347583" /translation="MELTGTDLWSVAIFTWIFILLVLMHRRQSWTSCYPGPVPWP VLGNLQVDLNNPYSYKIQNRVDGVSLOWAMPVWVINKMKAEVLLTCGEDTA DRPPVIFELHGFKPSQGMIFAPYPSWRQSRPESLSLRFGLGRKSLKEWIKAE GHLCDFTQAGQYINPNMLKATCNVIAELIFARRFEYEDPYILRMKLVESLGE LSGLPEVINTFPFLLHPLRLADKFLQSKSFIALVDNLTENRTFTWDPQSPRNLT AFLAEIKAGNPESFNENLMVVDLTFTAGLITTTSLWALLMLHPDVORR QOEIDVTQVHRPMPAQHMPYTNVIVHEVQRFQDVPVPLHPLRITSRDIEVDQFLI PKGTILLPNMSSVHMDDTWKEKPLRFHPHEFLDAQCHFPVKHEAFITFSAGRRSCLIGP
CDS	


```
Db 203 GTGGACCTGGATAAATCATGCTTGTACAGCTTGTAAAGCTTCAAAACCGCTATGGTGACGTG 262
Qy 69 PheSerLeuGlnLeuAlaTrpThrProValValValLeuAsnGlyLeuAlaValArg 88
Db 263 TTCAGCCTGCAGATGGCTGGAGCCTATGGTGTGATCAATGGACTGAAGCAATGAAG 322
Qy 89 GluAlaLeuValThrHisGlyGluAspThrAlaAspArgProProValProIleThrGln 108
Db 323 GAAGTGTCTGTACCTGTGGAGGAGACACTGTGTGACCGCCCTCAAGTGCCCACTTTGAG 382
Qy 109 IleLeuGlyPheGlyProArgSerGln 117
Db 383 TACCTGGGTGTGAAGCCTGGATCCCAAGGTGTGTCTTTCACCCCTACGGGCCCGAGTGG 442
Qy 117 ----- 117
Db 443 CGAGAGCAGAGGCGATCTCTGTGTCTACCTCGCGCAACTTTGGCCTGGGCAAGAAATCA 502
Qy 117 ----- 117
Db 503 CTGGAGGACTGGTGACCAAGGAGGCCAGACACTCTGTGTATGCTTACCGCCCGAGGCT 562
Qy 118 GlyArgProPheArgProAsnGlyLeuLeuAspLysAlaValSerAsnValIleAlaSer 137
Db 563 GCGCAGCCCATCAATCCCAACACCATGTGTGAACACGCTGTGTCAATGTGATTGCACT 622
Qy 138 LeuThrCysGlyArgArgPheGlyThrAspAspProArgPheLeuArgLeuLeuAspLeu 157
Db 623 CTCATTTTGGCCGTCGTCTTGAATATGAAGACCTTACCTCATCAGGATGAGAAAGTA 682
Qy 158 AlaGlnGluGlyLeuLysGluSerGlyPheLeuArgGluValLeuAsnAlaValPro 177
Db 683 CTGGAAGATAGTTTGACAGAAATCTCTGCTTAATCTCTGAGTCTTAAATGTTCCTCC 742
Qy 178 ValLeuLeuHisIleProAlaLeuAlaGlyLysValLeuArgPheGlnLysAlaPheLeu 197
Db 743 ATACTCTTCGCGATCCAGGACTGCTGCGAAGTCTTCCAAGCTCAGAAGTCCCTACTG 802
Qy 198 ThrGlnLeuAspGluLeuLeuThrGluHisArgMetThrTrpAspProAlaGlnProPro 217
Db 803 GCCATAGTGGAAATCTGTGTGATGAGATAGGAACACTGGGACCCCTGACCAGCCACCC 862
Qy 218 ArgAspLeuThrGluAlaPheLeuAlaGluMetGluLysAlaLysGlyAsnProGluSer 237
Db 863 CGAATTTGACTGATGCCTTCTTCGAGAGATAGAGAGTAAAGGGAATGCTCAGAGC 922
Qy 238 SerPheAsnAspGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetVal 257
Db 923 AGCTTCAATGATGAGAACCTGGCATGTTGTGTGTAGACCTGTTCACTGCAGGGATGTG 982
Qy 258 ThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGln 277
Db 983 ACCACTCAACACACTGTCCTGGGCCCTGCTGCTCATGATCCTGCATCCGATGTGCAG 1042
Qy 278 ArgArgValGlnGlnGluLeuAspValIleGlyGlnValArgArgProGluMetGly 297
Db 1043 CCAGAGTCCACAGAAATCATGTCGGTCAATAGGCGAGGTGCGGCATCCAGAGATGCA 1102
Qy 298 AspGlnAlaHisMetProThrThrAlaValIleHisGluValGlnArgPheGlyAsp 317
Db 1103 GACCAAGCTCGTATGCCCTTACACCAATGCTGTCTCATGAGGTACAGCGCTTTCGGGAC 1162
Qy 318 IleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPheArg 337
Db 1163 ATGTCTCCACTGAATTTGCGACGCATCAACAGTCGTGACATTAAGAGTGCGAGGACTCTC 1222
Qy 338 IleProLysGlyThrThrLeuIleThrAsnLeuSerSerValLeuLysAspGluAlaVal 357
Db 1223 ATCCCCAA----- 1230
Qy 358 TrpGluLysProPheArgPheHisProGluHisPheLeuAspAlaGlnGlyHisPheVal 377
Db 1230 ----- 1230
```

```
Qy 378 LysProGluAlaPheLeuProPheSerAlaGlyArgAlaCysLeuGlyGluProLeu 397
Db 1231 -----GGCCGAGATCATGCTGGGTGAGCCCTG 1260
Qy 398 AlaArgMetGluLeuPhePheThrSerLeuLeuGlnHisPheSerPheSerVal 417
Db 1261 GCCCGATGGAGCTCTCTCTTTCACGTGCTCCTGCGACACTTTAGCTTCTCAGTG 1320
Qy 418 ProThrGlyGlnProArgProSerHisGlyValPheAlaPheLeuValThrProSer 437
Db 1321 CCAATGACAGCCAGCCAGGCTAGAAACCTGTGTCTTTCCTTTCCGGTTGCCCTTAC 1380
Qy 438 ProThrGluLeuCysAlaValProArg 446
Db 1381 CCTTACCAGCTCTGTGTGTGTATGCGT 1407

RESULT 6
LOCUS BX422592 1201 bp mRNA linear EST 13-MAY-2003
DEFINITION BX422592 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CS0DM008Y007 5-PRIME, mRNA sequence.
ACCESSION BX422592
VERSION BX422592.1 GI:30659347
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10301.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgl-bin/cwuser.cgi?seq=CS0DM008AH04QPi&cluster=10301.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DM008AH04QPi.

FEATURES
location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM008Y007"
/tissue_type="fetal"
/dev_stage="FETAL LIVER"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Alignment Scores: 7.26e-125 Length: 1201
Pred. No.: 1246.50 Matches: 253
Score: 81.88% Conservative: 0
Best Local Similarity: 81.88% Mismatches: 5
Query Match: 53.50% Indels: 52
DB: 13 Gaps: 1

US-09-820-788A-2 (1-446) x BX422592 (1-1201)
Qy 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
Db -----
```

```
Db 94 ATGGGGCTAGAGCACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCTGCTCCTGGTG 153
Qy 21 AspLeuMethHisArgGlnAgtPAlaAlaArgTyrSerProGlyProLeuProLeu 40
Db 154 GACTGATGACCGCGCAACGCTGGCTGACGCTACTCACCAGGCCCTTGGCACATG 213
Qy 41 ProGlyLeuGlyAsnLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 214 CCGGGCTGGCAACCTGCTGCATGTGGACTTCARAAACACACATCTGCTTCGACCCAG 273
Qy 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
Db 274 TTGCGCGCGCTTGGGGAGCTGTTGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 333
Qy 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 334 CTCAATGGGCTGGCGCGTGGCGARGCGCTGGTGACCCACGCGGAGGACACCGCGGAC 393
Qy 101 ArgProProValProleThrGlnIleLeuGlyPheGlyProArgSerGln 117
Db 394 CGCCCGCTGTGCCCATCACCCAGATCTCTGGGTTTGGCGCGGCTTCCCAAGGGGTGTTTC 453
Qy 117 ----- 117
Db 454 CTGGCGGCTATGGCGCGCGTGGCGGAGCAGAGCGGCTTCTCGTCTCCACCTTGGCG 513
Qy 117 ----- 117
Db 514 AACTTGGGCTGGCAAGAGTCTGGAGCAGTGGTGTGACGAGGAGGCGGCTGCCCTT 573
Qy 118 ----- GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
Db 574 TGTGCGCGCTTGGCAACCACTCGGAGCGCCCTTTCGCCCAACGCTCTTTGGACAAA 633
Qy 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149
Db 634 GCGTGAGCAGTGTATGCTCCCTCACCCTGGCGCGCGCTTCGAGTACGACGACCTT 693
Qy 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
Db 694 CGCTTCTCAGGCTGTGGACCTAGCTCAGGAGGAGTGAAGGAGGAGTGGGCTTCTG 753
Qy 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
Db 754 CGCGAGGTGTGAATGTGTCCCGCTCTCTGCATATCCACGCGTGGCTGGCAAGGTC 813
Qy 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
Db 814 CTACGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGCTGCTAACTGAGCAGGATG 873
Qy 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 874 ACCTGGGACCCAGCCAGCCCGCCGAGACCTGACTGAGGCTTCTTGGCAGAGATGGAG 933
Qy 230 LysAlaLysGlyAsnProGluSerPheAsnAspGluAsnLeuArgIleValValAla 249
Db 934 AAGSCAAGGGGAACCTTGAAGAGCTTCAATGATGAGAACTG-CGCATAGTGTGGTCT 992
Qy 250 AspLeuPheSerAlaGlyMetValThr 258
Db 993 GACCTGTCTCTCCCGGGATGGTGACA 1019
```

```
RESULT 7
CK032845/c
LOCUS
DEFINITION
AGENCOURT_15196984 NIH_MGC_195 Homo sapiens cDNA clone
IMAGE:7002156 5', mRNA sequence.
ACCESSION
VERSION
CK032845
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE

1 (bases 1 to 814)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Narayan Bhat
CDNA Library Preparation: Bhat Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: IRBKA row: C column: 10
High quality sequence start: 10
High quality sequence stop: 718.

FEATURES

source

1..814
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7002156"
/tissue_type="mixed"
/lab_host="DH5A (TI phage-resistant)"
/clone_lib="NIH MGC 195"
/note="Vector: pDNR-Dual; Site 1: loxP-Sall; Site 2:
loxP-HindIII; Clones from this library have been
PCR-amplified using gene-specific primers to contain the
complete open reading frame (based on known gene sequences
available from NCBI's RefSeq). Template for PCR is cDNA
derived from either pooled cytoplasmic polyA RNA from 30
cells lines or pooled total RNA from 10 different tissues
(from BD Biosciences/Clontech and Washington University).
PCR products are directionally cloned into the loxP sites
of the pDNR-Dual vector. Library constructed by Dr.
Narayan Bhat, Earl Bere III and Hongling Liao (Gene
Expression Laboratory, Research Technology Program, SAIC
Frederick, NCI-Frederick, Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK.prsv.dat
a Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 3.56e-123 Length: 814
Score: 1228.50 Matches: 248
Percent Similarity: 96.89% Conservative: 1
Best Local Similarity: 96.50% Mismatches: 7
Query Match: 52.73% Indels: 5
DB: 14 Gaps: 1
US-09-820-788A-2 (1-446) x CK032845 (1-814)

Qy 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuThrGluHisArgMet 209
Db 799 CTACGCTTCCAA---AGCTTCTCTCCCGGAGCT-GATGAGTG-CTAACTGAGCAGAGTG- 746
Qy 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 745 ACCTGGAC-CCAGCCCGAGCCCCCGAGACCTGACTGAGGCTTCTTGGCAGAGATGGAG 687
Qy 230 LysAlaLysGlyAsnProGluSerPheAsnAspGluAsnLeuArgIleValValAla 249
Db 686 AAGSCAAGGGGAACCTTGAAGAGCTTCAATGATGAGAACTGAGAACTGAGTGTGGTCT 627
Qy 250 AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTrpGlyLeuLeuLeu 269
Db 626 GACCTGTCTCTCCCGGGATGGTGACCACTGAGCAGCGCTGGCGGCTCTCTGCTC 567

QY 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGlnLeuAspValIleGly 289
 Db 566 ATGATCTCATCTCGGATGTGAGGCGCGTGTCCAAAGAGAGATGAGACGACGTGATAGGG 507
 QY 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyThrThrAlaValIle 309
 Db 506 CAGGTGGCGGACAGAGATGGTGACAGGCTCATGTCCTTACCACTCCGCTGATT 447
 QY 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
 Db 446 CATGAGTGCAGCGCTTTGGGACATCGTCCCTGGGTGTGACCATATGACATCCCGT 387
 QY 330 AspIleGluValGlnGlyPheArgIleProGlyThrThrLeuIleThrAsnLeuSer 349
 Db 386 GATCTGAAGTACAGGCTTCGCGCATCCCTAAGGAAGACACTCATCAACCACTGTCA 327
 QY 350 SerValLeuLeuAspGluAlaValTrpGluLeuProPheArgPheHisProGluHisPhe 369
 Db 326 TCGGTGTGAAGATGAGCGCTGTGGGAAGACCTTCGCTTCCACCCGACACTTC 267
 QY 370 LeuAspAlaGlnGlyHisPheValIleProGluAlaPheLeuProPheSerAlaGlyArg 389
 Db 266 CTGGATGCCAGGCGCACTTTGTGAAGCGGAGGCTTCCTGCTTCTCAGCAGCGCGC 207
 QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrSerLeu 409
 Db 206 CTGTGATGCTCGGGAGCGCTGCGCGCATGGAGCTTCTCTTCTTCACTCCCTG 147
 QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
 Db 146 CTGAGCACTTCAGCTTCTCGTGCCACTGAGAGCGCGCGCCAGCCACCATGCTGTC 87
 QY 430 PheAlaPheLeuValThrProSerProTyThrGluLeuCysAlaValProArg 446
 Db 86 TTTGCTTTCGTGTCACCCATCCCTCATGAGCTTTGTGCTGTGCCCGC 36

RESULT 8

CF271947 798 bp mRNA linear EST 19-NOV-2003
 LOCUS AGENCOURT 15197016 NIH MGC 195 Homo sapiens cdna clone
 DEFINITION IMAGE:7002157 5', mRNA sequence.

ACCESSION CF271947
 VERSION 1
 KEYWORDS EST.

SOURCE CF271947.2 GI:38453389
 Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 798)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT On Aug 12, 2003 this sequence version replaced gi:33627859.

Contact: Daniela S. Gerhard, Ph.D.,
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Narayan Bhat
 cDNA Library Preparation: Bhat Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: IRBK4 row: c column: 11

High quality sequence start: 8

High quality sequence stop: 748.

Location/Qualifiers

1..798

/organism="Homo sapiens"

/mol_type="mRNA"

FEATURES

source

/db_xref="taxon:9606"
 /clone="IMAGE:7002157"
 /tissue_type="mixed"
 /lab_host="DH5A (Ti phage-resistant)"
 /clone_lib="NIH MGC 195"

/note="Vector: pDNR-Dual; Site 1: loxp-SalI; Site 2:
 loxp-HindIII; Clones from this library have been
 PCR-amplified using gene-specific primers to contain the
 complete open reading frame (based on known gene sequences
 available from NCBI's RefSeq). Template for PCR is cDNA
 derived from either pooled cytoplasmic polyA RNA from 30
 cells lines or pooled total RNA from 10 different tissues
 (from BD Biosciences/Clontech and Washington University).
 PCR products are directionally cloned into the loxp sites
 of the pDNR-Dual vector. Library constructed by Dr.
 Narayan Bhat, Earl Bere III and Hongling Liao (Gene
 Expression Laboratory, Research Technology Program, SAIC
 Frederick, NCI-Frederick, Frederick, MD 21702). For
 information on which gene each clone represents, please
 visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK.presv.dat
 a Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 3,91e-123 Length: 798
 Score: 1228.00 Matches: 241
 Percent Similarity: 97.98% Conservative: 2
 Best Local Similarity: 97.18% Mismatches: 3
 Query Match: 52.70% Indels: 4
 DB: 14 Gaps: 1

US-09-820-788A-2 (1-446) x CF271947 (1-798)

QY 201 AspGluLeuLeu-----ThrGluHisArgMetThrTrpAspProAlaGlnProProArg 218
 Db 778 GACCATTTGATGAGTGTCTACTGAGCAGCAG-ATGACNTGGAC-CCAGCCCGCCCCCGA 721
 QY 219 AspLeuThrGluAlaPheLeuAlaGluMetGluLeuAlaGlyAsnProGluSerSer 238
 Db 720 GACCTGACTGAGGCGCTTCTGGCAGAGATGAGAAGGCCAAGGGGAAACCTTGAGCAGC 661
 QY 239 PheAsnAspGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetValThr 258
 Db 660 TTCAATGATGAGAACCTGTGTCATAGTGTGCTGACCTGTTCTTGCCTGGGATGGTGACC 601
 QY 259 ThrSerThrThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGlnArg 278
 Db 600 ACCTCGACACGCTGGCTGGGCGCTCCTGTCTCATGATCCTACATCCGGATGTGCAGGCG 541
 QY 279 ArgValGlnGlnGluLeuAspValIleGlyGlnValArgArgProGluMetGlyAsp 298
 Db 540 CGTGTCCAACAGGAGATCGACGCTGATAGGGCAGGTGCGCGCAGCAGAGATGGGTGAC 481
 QY 299 GlnAlaHisMetProTyThrThrAlaValIleHisGluValGlnArgPheGlyAspIle 318
 Db 480 CAGGCTCACATGCCCTCACACCCTGATTATGAGGTGCGAGCGCTTTGGGGACATC 421
 QY 319 ValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPheArgIle 338
 Db 420 GTCCCCCTGGGTGTGACCCATATGACATCCGTGACATCGAAGTACAGGGCTTCCGCATC 361
 QY 339 ProLysGlyThrThrLeuIleThrAsnLeuSerSerValLeuLysAspGluAlaValTrp 358
 Db 360 CCTAAGGGAACGACACTCATCCCAACCTGTCATCGGTGCTGGAAGGATGAGCGCGTCTGG 301
 QY 359 GluLysProPheArgPheHisProGluHisPheLeuAspAlaGlnGlyHisPheValIys 378
 Db 300 GAGAAGCCCTTCCGCTTCCACCCGAAACACTTCTCGATGCCAGGGGCACATTGTGTGAAG 241
 QY 379 ProGluAlaPheLeuProPheSerAlaGlyArgArgAlaCysLeuGlyGluProLeuAla 398
 Db 240 CCGGAGGCGCTTCTGCGCTTCTTCAGCAGCGCGCGCTGCATGCTCGGGAGCCCTCGGCC 181

QY 399 ArgMetGluLeuPheLeuPheThrSerLeuLeuGlnHisPheSerPheSerValPro 418
Db 180 CGCATGAGCTCTTCTCTTACCTCCCTGCTGAGCACTTCAGCTTCTCGGTGCC 121
QY 419 ThrGlyGlnProArgProSerHisGlyValPheAlaPheLeuValThrProSerPro 438
Db 120 ACTGGACAGCCCGCCGAGCACCACCATGGTGTCTTTCTGCTGAGCCCATCCCC 61

QY 439 TyrGluLeuCysAlaValProArg 446

Db 60 TATGACCTTGTGCTGTGCCCGC 37

RESULT 9

CK032848/c

LOCUS

DEFINITION

AGENCOURT 15196888 NIH MGC 195 Homo sapiens cDNA clone

IMAGE:7002153 5', mRNA_sequence.

CK032848

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cga@nci.nih.gov

Tissue Procurement: Narayan Bhat

cDNA Library Preparation: Bhat Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: IRBK4 row: c column: 07

High quality sequence start: 12

High quality sequence stop: 726.

Location/Qualifiers

1..803

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:7002153"

/tissue_type="mixed"

/lab_host="DH5A (T1 phage-resistant)"

/clone_lib="NIH MGC 195"

/note="Vector: pDNR-Dual; Site 1: loxp-SalI; Site 2:

loxP-HindIII; Clones from this library have been

PCR-amplified using gene-specific primers to contain the

complete open reading frame (based on known gene sequences

available from NCBI's RefSeq). Template for PCR is cDNA

derived from either pooled cytoplasmic polyA RNA from 30

cells lines or pooled total RNA from 10 different tissues

(from BD Biosciences/Clontech and Washington University).

PCR products are directionally cloned into the loxp sites

of the pDNR-Dual vector. Library constructed by Dr.

Narayan Bhat, Earl Bere III and Hongling Liao (Gene

Expression Laboratory, Research Technology Program, SAIC

Frederick, NCI-Frederick, Frederick, MD 21702). For

information on which gene each clone represents, please

visit our anonymous ftp site at

ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat

a Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 4,48e-123 Length: 803
Score: 1227.50 Matches: 244
Percent Similarity: 96.84% Conservative: 1
Best Local Similarity: 96.44% Mismatches: 8
Query Match: 52.68% Indels: 3
DB: 14 Gaps: 0

US-09-820-788A-2 (1-446) x CK032848 (1-803)

QY 194 LysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMetThrTrpAspPro 213

Db 792 AAGATTTTCTTGACCCAGCT-GATGAGTG-GTAAATGAGCAGAGATGACNGG-GACCCA 736

QY 214 AlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGluLysAlaLysGly 233

Db 735 GCCCAGCCCCCGAGACTTGACTGAGCGCTTCCCGCAGAGATGAGAGAGCCCAAGGG 676

QY 234 AsnProGluSerSerPheAsnAspGluLeuArgLleValValAlaAlaPheLeuPheSer 253

Db 675 AACCTGTAGAGCAGCTTCAATGATGAGAACCTGTGATAGTGGCTGACCTGTCTCT 616

QY 254 AlaGlyMetValThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 273

Db 615 GCCGGATGGTGACCCACCTCGACACGCTGGCGCTGGGGCTCTCTGCTCATGATCTTACAT 556

QY 274 ProAspValGlnArgArgValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 293

Db 555 CCGGATGTGACGCGCGTGTCCCAACAGGAGATCGACGCGTGTATAGGCGAGTGGCGCA 496

QY 294 ProGluMetGlyAspGlnAlaHisMetProTyThrThrAlaValLleHisGluValGln 313

Db 495 CCAGAGATGGGTGACCGAGTTTCCATGATGAGTGGCTGATTCATGAGGTGCGAG 436

QY 314 ArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluVal 333

Db 435 CCCTTTGGGACATCGTCCCTGGGTGGACCATATGACATCCCGTGCATTCGAGTGA 376

QY 334 GlnGlyPheArgIleProLysGlyThrThrLeuLleThrAsnLeuSerSerValLeuLys 353

Db 375 CAGGCTTCCGATCCCTAAGGGAGACACACTCATCCCAACCTGTCTCGTGTCTGAAG 316

QY 354 AspGluAlaValTrpGluLysProPheArgPheHisProGluHisPheLeuAspAlaGln 373

Db 315 GATGAGGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCCGAACACTTCTCGATGCC 256

QY 374 GlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArgAlaCysLeu 393

Db 255 GGCCACTTTGTGAAGCGGAGCGCTTCTGCTTCTTCTGAGAGCGCGCGTGCATCCCTC 196

QY 394 GlyGluProLeuAlaArgMetGluLeuPhePhePheThrSerLeuLeuGlnHisPhe 413

Db 195 GGGGAGCCCTGCGCCGCGATGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 136

QY 414 SerPheSerValProThrGlyGlnProArgProSerHisHisGlyValPheAlaPheLeu 433

Db 135 AGCTTCTGCTGCTGCCACTGGACAGCCCGCGCCAGCCACCATGGTGTCTTTGCTTCT 76

QY 434 ValThrProSerProTyThrGluLeuCysAlaValProArg 446

Db 75 GTGACCCCATCCCTTATGAGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 37

RESULT 10

AK004933 3079 bp mRNA linear HTC 20-SEP-2003

LOCUS Mus musculus adult male liver cDNA, RIKEN full-length enriched

DEFINITION library, clone:130007K12 product:similar to CYTOCHROME P450 2D3

(EC 1.14.14.1) (CYP1D3) (P450-DB3) (DEBRISOQUINE 4-HYDROXYLASE)

(Rattus norvegicus), full insert sequence.

ACCESSION AK004933

VERSION AK004933.1 GI:12836492

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

QY 167 GlyPheLeuArgGluValLeuAenAlaValProValLeuLeuHisIleProAlaLeuAla 186
Db 563 GGTCTGTTTCCAGAGGTTCTGAACACATTTCTCCATGACATCCAGGGTTGGCT 622
QY 187 GlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuThrGlu 206
Db 623 GACAAAGTCTTCCCTGGGCAAGACATTTCTCACCTGGTAAATAGAGCTGGTACTGAG 682
QY 207 HisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAla 226
Db 683 CACAAGAGGACCTGGGACCTTGACACACCTCGTGACCTGACTGATGCTTCTGCT 742
QY 227 GluMetGluLysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIle 246
Db 743 GAAATGGAGAAGCAAGGGGAAACCTTAAGAGTAGCTTCAATGAAGCAAACTACGCTCG 802
QY 247 ValValAlaAspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGly 266
Db 803 GTTGTCTTTGACCTGTTGGTGTGGATTTGTGACAGTTCAATCACTCTGACCTGGGCC 862
QY 267 LeuLeuLeuMetIleLeuHisProAspValGlnArgValGlnGlnGlnIleAspAsp 286
Db 863 CTGCTGCTCATGATCTTGCATCCAGATGTGCAGCGCGAGTGCAGAGGAAATCGACGAG 922
QY 287 ValIleGlyGlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThr 306
Db 923 GTCATAGGCGAGTGGGTGTCCAGAGATGGCAGACCGGCCACATGCGCTACACCAAT 982
QY 307 AlaValIleHisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMet 326
Db 983 GCTGTCAATCATGAGTGGAGCGCTTTGCAGACATTTGTCACCAATGTCCTCCACACAAG 1042
QY 327 ThrSerArgAspIle-GluValGlnGlyPheArgIleProLysGlyThrThrLeuIleTh 346
Db 1043 ACTTCTCATGATTTGGAGTAGACGGCTTCTTATCCCTAAGGGACGACCTCATCC 1102
QY 346 rAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisPr 366
Db 1103 CAACCTGTCTCGCGCTAAAGATGAGACTGTCTGGGAGAAGCCCTCCGCTTCATCC 1162
QY 366 oGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSe 386
Db 1163 TGAACACTTCTGGATGCCAGGCGCAATTTGTGAAGCCTGAGGCTTCATGCCATTCTC 1222
QY 386 rAla----- 387
Db 1223 AGC-AGGTGCTTGGGTGCTGGTCCCTGCTCTCTCCCGGGTGCCTTGGAGGTGG 1281
QY 387 ----- 387
Db 1282 AGCACTGTGGGGCATCAGGCTTGGGGAATAAAGGGGTGGGAGAACCCGACTGGAC 1341
QY 387 ----- 387
Db 1342 ATGGGGGACTGCTTGCATCTCTATGCGCGCCGGGTGAGCGTCTGGCTTGGGAAGCTA 1401
QY 387 ----- 387
Db 1402 TGAACTCAGTTCAAGGGGGGCGAGAGAGAACGAGTCAGGGTCTCTGGGGCTCATGGA 1461
QY 387 ----- 387
Db 1462 GGCCAGAGACGAGGTTCTCAGTCTCACACATCCAGATGTCACATGTATGTGCGTGAAG 1521
QY 387 ----- 387
Db 1522 AGCTGAGAACAGGCCCTCAGATGGACTCTGCGCGGGGCGGAGGAGGGTAAGGGGAGA 1581
QY 387 ----- 387
Db 1582 GCCCTGGCAGTTTACTGCGAGGATGAGAGAACTGACTAGTTTACGGCTGCTAGCTCGGTG 1641
QY 387 ----- 387

Db 1642 GGTGCTTGGTGGTGGCCGCTGAGTGCAGAGAGCCCTTCCAGTGGAGGT 1701
QY 387 ----- 387
Db 1702 TAGATGTGGCTCTTTAGAGGAAGCCCTACTCTGTTTCCCAACACAGTGAAGTCTCAATGA 1761
QY 387 ----- 387
Db 1762 GCAGAGATTGTCTATGTTTGTAGAGCTTTATTGTTAGAAAAACAGGGAGAAAGAGAGAGG 1821
QY 387 ----- 387
Db 1822 TAGAAAGAAATAGCCAGCCATAGCCAGTGGAGAAAAACCGAGAGGGGCTGAAAG 1881
QY 387 ----- 387
Db 1882 AGAGCTAGAGATAAGATTAAAGAGACTGTGGGAGCTTTAAGAGAACAAAGAGGGGCCAAGCA 1941
QY 387 ----- 387
Db 1942 GCCCCTTTTATAGTGGCTGGCTATCTTGCAGTTTGCAGGGCAGTTGTGTGAAGAGCAT 2001
QY 387 ----- 387
Db 2002 ACCTGGCTATAGCCAGATACTGTGGGATGGAGTCTAGCCAGATAACAGGAGTTTGGGG 2061
QY 387 ----- 387
Db 2062 ATGTGACTGGTAGTCACAGGATTATGGAGCTGGGGGCTCCGTGGTGTCCAGGACCATATGT 2121
QY 387 ----- 387
Db 2122 CTGTTTGGGAGATGGCTCACTGCTCTGTAGAGTTTCTACTGAGTCTCCAGAGTAAGTCT 2181
QY 387 ----- 387
Db 2182 CACCTGACCAGAAAAAGGCTGCTTTCACAGTCCCAAAAGCACCATCAAGTCCAGCA 2241
QY 388 -----GlyArgArgAlaCysLeuGlyGluProLeuAlaAr 399
Db 2242 TTGACTGAGCCCTTTCTGACCCAGAGTGTCTGAGCATGCTGGGGAGCCCCCTGGCCCG 2301
QY 399 gMetGluLeuPheLeuPheThrSerLeuLeuGlnHisPheSerPheSerValProTh 419
Db 2302 CATGAGGCTCTCTCTTCTTCTCCTGCTCTCTGAGCGCTTACCTTCTTAGTGCTGC 2361
QY 419 rGlyGlnProArgProSerHisHisGlyValPheAlaPheLeuValThrProSerProTy 439
Db 2362 TGGACAGCCCGACCCAGCCAGTACTATGGCATCTTTATACATTTCTGTTAGCCCCCTC 2421
QY 439 rGluLeuCysAlaValProArg 446
Db 2422 TCAGCTCTGTGCATTCACACGT 2443

RESULT 11
CK032844/c

LOCUS

DEFINITION AGNCOURT_15197048 NIH_MGC_195 Homo sapiens cDNA clone
IMAGE:7002158 5', mRNA sequence.

ACCESSION

CK032844

VERSION

CK032844.1

GI:38558768

EST.

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 767)

NIH-MGC <http://mgs.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

CK032844 767 bp mRNA linear EST 26-NOV-2003

AGNCOURT_15197048 NIH_MGC_195 Homo sapiens cDNA clone

IMAGE:7002158 5', mRNA sequence.

CK032844

CK032844.1

GI:38558768

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 767)

NIH-MGC <http://mgs.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9apbs-x@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: Bhat Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: IRBK4 row: c column: 12
High quality sequence start: 5
High quality sequence stop: 714.
Location/Qualifiers
1. .767
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7002158"
/tissue_type="mixed"
/lab_host="DH5A (T1 phage-resistant)"
/clone_lib="NIH MGC 195"
/note="Vector: pDNR-Dual; Site 1: loxP-Sall; Site_2:
loxP-HindIII; Clones from this library have been
PCR-amplified using gene-specific primers to contain the
complete open reading frame (based on known gene sequences
available from NCBI's RefSeq). Template for PCR is cDNA
derived from either pooled cytoplasmic polyA RNA from 30
cells lines or pooled total RNA from 10 different tissues
(from BD Biosciences/Clontech and Washington University).
PCR products are directionally cloned into the loxP sites
of the pDNR-Dual vector. Library constructed by Dr.
Narayan Bhat, Earl Bere III and Hongling Liao (Gene
Expression Laboratory, Research Technology Program, SAIC
Frederick, NCI-Frederick, Frederick, MD 21702). For
information on which gene each clone represents, please
visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK.presv.dat
a Note: this is a NIH MGC Library."

ORIGIN

Alignment Scores:					
Pred. No.:	1,29e-122	Length:	767		
Score:	1223.00	Matches:	237		
Percent Similarity:	98.34%	Conservative:	0		
Best Local Similarity:	98.34%	Mismatches:	4		
Query Match:	52.49%	Indels:	1		
DB:	14	Gaps:	0		
US-09-820-788A-2 (1-446) x CK032844 (1-767)					
Qy	206	GluHisArgMetThrTyrAspProAlaGlnProProArgAspLeuThrGluAlaPheLeu	225		
Db	756	GAGCACAGGATGACCTGGAC - CCAGCCCGCCCCCGAGACCTGTGANTGAGGCCTTCCTG	698		
Qy	226	AlaGluMetGluLysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArg	245		
Db	697	GCAGAGNTGAGAAGGCCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGACCTGTGC.	638		
Qy	246	IleValValAlaAspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrp	265		
Db	637	ATAGTGGTGCTGACCCTGTTCTCTCGCGGATGGTGACCACCTCGACCACCGCTGGCGCTGG	578		
Qy	266	GlyLeuLeuLeuMetIleLeuHisProAspValGlnArgArgValGlnGlnGluLeuAsp	285		
Db	577	GGCGCTCTGCTCATGATCCTACATCCGGATGTGCAGCGCGCTGTCTCAACAGGAGATCGAC	518		
Qy	286	AspValIleGlyGlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThr	305		
Db	517	GAGGTGATAGGGCAGGTGGCGGCACAGAGATGGGTGACCAGGCTCATGCCCCTACACC	458		
Qy	306	ThrAlaValIleHisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHis	325		

Db	457	ACTCCCTGATTCATGAGGTGCACGCCCTTTGGGACATCGTCCCTCGGTGTGACCCCAT	398
Qy	326	MetThrSerArgAspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIle	345
Db	397	ATGACATCCCGTGACATCGAAGTACAGGGCTTCGCATCCCTAAGGGAACGACACTCATC	338
Qy	346	ThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHis	365
Db	337	ACCAACCTGTTCATCGGTGCTGAAGAGATGAGGCCCTCTGGGAGAAAGCCCTTCGCGCTTCAC	278
Qy	366	ProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPhe	385
Db	277	CCCGAACACTTCCTCGGATGCCCGAGGCCACTTTGTGAAGCCGAGGCCCTTCCTGCCTTC	218
Qy	386	SerAlaGlyArgArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhe	405
Db	217	TCAGCAGCGCGCGTGATCGCTCGGGAGGCCCTTCGGCCGCATGGAGCTCTTCCTCTTC	158
Qy	406	PheThrSerLeuLeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSer	425
Db	157	TTCACTCCCTGTCTGCAGCATTCTCAGCTTCTCGGTGGCCACTGGACAGCCCGGCCAGC	98
Qy	426	HisHisGlyValPheAlaPheLeuValThrProSerProTyrgluLeuCysAlaValPro	445
Db	97	CACCATGGTGTCCTTTGCTTCTTCCTGGTACCCCATCCCGCCGATGAGCTTTGTGTGTGCC	38
Qy	446	Arg 446	
Db	37	CGC 35	
RESULT 12	CK032843/c		
LOCUS	AGENCOURT 15196697 NIH MGC 195 Homo sapiens cDNA clone	780 bp mRNA linear EST 26-NOV-2003	
DEFINITION	IMAGE:7002159 5', mRNA sequence.		
ACCESSION	CK032843		
VERSION	CK032843.1	GI:38558767	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 780)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgaabs-r@mail.nih.gov Tissue Procurement: Narayan Bhat cDNA Library Preparation: Bhat Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: IRBK4 row: d column: 01 High quality sequence start: 10 High quality sequence stop: 719. Location/Qualifiers 1..780		
FEATURES	source		

source

complete open reading frame (based on known gene sequences available from NCBI's RefSeq). Template for PCR is cDNA derived from either pooled cytoplasmic polyA RNA from 30 cells lines or pooled total RNA from 10 different tissues (from BD Biosciences/Clontech and Washington University). PCR products are directionally cloned into the loxp sites of the pDNR-Dual vector. Library constructed by Dr. Narayan Bhat, Earl Bere III and Hongling Liao (Gene Expression Laboratory, Research Technology Program, SAIC Frederick, NCI-Frederick, Frederick, MD 21702). For information on which gene each clone represents, please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
a Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 1,93e-122 Length: 780
Score: 1221.50 Matches: 241
Percent Similarity: 97.97% Conservative: 0
Best Local Similarity: 97.97% Mismatches: 5
Query Match: 52.42% Indels: 3
DB: 14 Gaps: 0

US-09-820-788A-2 (1-446) x CK032843 (1-780)

QY 201 AspGluLeuLeuThrGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeu 220
DB 770 GATGAGCTGCTAAT-GAGCAGAC-ATGACCTGGAC-CCAGCCAGACCCCCCGAGACCTG 714
QY 221 ThrGluAlaPheLeuAlaGluMetGluLysAlaLysGlyAsnProGluSerSerPheAsn 240
DB 713 ACTGAGCGCTTCCCGGAGAGATGGAGAGCCAGCGGAACCTCGAGACGACTTCAAT 654
QY 241 AspGluAsnLeuArgileValValAlaAspLeuPheSerAlaGlyMetValThrThrSer 260
DB 653 GATGAGAACCTGTGCATAGTGGTGGCTGACCTTCTCTCGCGGAGTGTGACCACTCG 594
QY 261 ThrThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGlnArgArgVal 280
DB 593 ACCACGCTGGCTGGGGCTCTGCTCATGATCTTACATCCGATGTGAGCCCGCTGTC 534
QY 281 GlnGlnGluIleAspValIleGlyGlnValArgArgProGluMetGlyAspGlnAla 300
DB 533 CACAGGAGATGACGACGATGATAGGCGAGTGGCGGACACGAGATGGTGACAGGCT 474
QY 301 HisMetProTyThrThrAlaValIleHisGluValGlnArgPheGlyAspIleValPro 320
DB 473 CACATGCCCTTACACCACTGCCGTGATTCATGAGGTGACGCGCTTGGGACATCGTCCCC 414
QY 321 LeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPheArgIleProLys 340
DB 413 CTGGGTGTGACCATATGATCCCTGTGATCCAGTACAGGCTTCCGATCCCTAAG 354
QY 341 GlyThrThrLeuIleThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLys 360
DB 353 GGAACGACACTCATCCAACTGTTCATCGTCTGAAGGATGAGCGCTCTGGGAGAAG 294
QY 361 ProPheArgPheHisProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGlu 380
DB 293 CCGTTCGGTTCACCCCGAACACTTCTCGTGTGACCGGCGCACTTGTGAAGCGGAG 234
QY 381 AlaPheLeuProPheSerAlaGlyArgArgAlaCysLeuGlyGluProLeuAlaArgMet 400
DB 233 GCCTTCTGCTTCTCTCAGAGGCGCCCGTGCATGCTCGGGAGCCCTTGGCCCGCATG 174
QY 401 GluLeuPheLeuPhePheThrSerLeuLeuGlnHisPheSerPheSerValProThrGly 420
DB 173 GAGCTCTTCTCTTCTTACCTCCCTGCTGCAGCACTTCAGCTTCTCGTGGCCCACTGA 114
QY 421 GlnProArgProSerHisHisGlyValPheAlaPheLeuValThrProSerProTyGlu 440
DB 113 CAGCCCGCGCCAGCACCACCATGTGTCTTTTGTCTTCTGTTGACCCCACTCCCGCTATGAG 54

QY 441 LeuCyEAlaValProArg 446
DB 53 CTTTGTGCTGTGCCCCGC 36

RESULT 13
CK032846/c

LOCUS
DEFINITION
IMAGE: 7002155 5', mRNA-sequence.
ACCESSION
CK032846
VERSION
CK032846.1 GI:38558770
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT

CK032846 804 bp mRNA linear EST 26-NOV-2003
AGENCOURT 15196952 NIH_MGC 195 Homo sapiens cDNA clone
IMAGE: 7002155 5', mRNA-sequence.
CK032846
CK032846.1 GI:38558770
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 804)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: Bhat Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: IRBK4 row: C column: 09
High quality sequence start: 12
High quality sequence stop: 686.
Location/Qualifiers

FEATURES

source

1..804

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7002155"
/tissue_type="mixed"
/lab_host="DH5A (T1 phage-resistant)"
/clone_lib="NIH_MGC 195"
/note="Vector: pDNR-Dual; Site_1: loxp-Sall; Site_2: loxp-HindIII; Clones from this library have been PCR-amplified using gene-specific primers to contain the complete open reading frame (based on known gene sequences available from NCBI's RefSeq). Template for PCR is cDNA derived from either pooled cytoplasmic polyA RNA from 30 cells lines or pooled total RNA from 10 different tissues (from BD Biosciences/Clontech and Washington University). PCR products are directionally cloned into the loxp sites of the pDNR-Dual vector. Library constructed by Dr. Narayan Bhat, Earl Bere III and Hongling Liao (Gene Expression Laboratory, Research Technology Program, SAIC Frederick, NCI-Frederick, Frederick, MD 21702). For information on which gene each clone represents, please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
a Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 1.34e-121 Length: 804
Score: 1214.00 Matches: 241
Percent Similarity: 95.26% Conservative: 0
Best Local Similarity: 95.26% Mismatches: 11
Query Match: 52.10% Indels: 3
DB: 14 Gaps: 1

US-09-820-788A-2 (1-446) x CK032846 (1-804)

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10301.r For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAH018ZE100P1sclcluster=10301.r>. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Paradise Avenue Genoscope sequence ID : CS0BAH018ZE100P1

FEATURES

1. 1088
 Location/Qualifiers
 Faraday Avenue Genoscope sequence

/new_stages="Homo sapiens FETAL LIVER"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMWSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer; five primed cDNA
enriched, double-strand cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMWSPORT 6
vector. Library was not normalized."

ORIGIN

Alignment Scores:		
Pred. No.:	1.13e-120	Length:
Score:	1207.50	Matches:
Percent Similarity:	80.24%	Conservative:
Best Local Similarity:	79.94%	Mismatches:
Query Match:	51.82%	Indels:
DB:	13	Gaps:
		2
		58
		13
		263
		1088

US-09-820-788A-2 (1-446) x BX432878 (1-1088)

Qy	1	MetGlyLeuGluAlaLeuValProLeuAlaValIleValIlePheLeuLeuVal	20
Db	67	ATGGGGCTAGAAAGCACTGGTGGCCCTGATAGTGGCCATCTTCCTGCTCGTGGT	126
Qy	21	AspLeuMetHisArgArgGlnArgTyrAlaAlaArgTyrSerProGlyProLeuProLeu	40
Db	127	GACCTGATGCAACCGGCGCAACCGCTGGGCTGCACGCTACTCACCGAGCCCTTGCACCTG	186
Qy	41	ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln	60
Db	187	CCGGCTGGGGCAACTGCTGCAATGTGACATTCAGAACACACCATACTGCTTCGACCAG	246
Qy	61	LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaThrThrProValValVal	80
Db	247	TTGGGCGCGCTTCGGGAGCGTGTTCACGCTGCAGCTGGCGTGCACGCGGCTGTCGTG	306
Qy	81	LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp	100
Db	307	CTCAATGGGCTGGCGCGCTGGCGAGCGCTGTGTGACCCACCGCGAGGACACCGCCGAC	366
Qy	101	ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln	117
Db	367	CGCCCGCTGTGCCCATCACCCAGATCTGGGTTCGGCGCGTTCACAGGGGTGTTTC	426
Qy	117	-----	117
Db	427	CTGGCGCGCTATGGGCGCGCTGGCGCGAGCAGAGGCGCTTCTYCGTCTCCACTTCGCGC	486
Qy	117	-----	117
Db	487	AACTGGGCGCTGGGCAGAGATCGCTGTGGAGCAGTGGGTGACCGAGGAGCGCGCTGCCTT	546
Qy	118	-----GlyArgProPheArgProAsnGlyLeuLeuAspLys	129
Db	547	TGTGGCGCCTTCGCCAACCATCTCGAGCGCCCTTCGCCCCACCGTCTCTTGGACAAA	606
Qy	130	AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluThrAspAspPro	149
Db	607	GCGGTGAGCAACGTGATCGCTCCCTCACTCGCGGGCGCGCTTCGAGTAGACGACCCCT	666

Qy	194	LysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMetThrTrpAspPro	213
Db	790	AAAGGCTTCCTGACCAGCTGG---AGAGCTGCTACTGAGCACAGGATGACNTGGAC-CCA	735
Qy	214	AlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGluLysAlaLysGly	233
Db	734	GCCACGCCCCCGGA-GACCTGAATGAGGCCCTCTCGCAGAGATGGGAAGGCCCAAGGGG	676
Qy	234	AsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAlaAspLeuPheSer	253
Db	675	AACCTTGAGAGGAGCTTCATGATGAGAAACCTGTGCATGTGTGGTGACCTGTCTCT	616
Qy	254	AlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHis	273
Db	615	GCCGGGATGGTGACCACCTCGACACAGCTGGCTGGGCCCTCTCTGCTCATGATCCTACAT	556
Qy	274	ProAspValGlnArgArgValGlnGlnGlnIleAspAspValIleGlyGlnValArgArg	293
Db	555	CCGGATGTCCAGCGCCGVTCCACAGAGAGATCCACCGTGTATAGGGCAGGTGCGCGCA	496
Qy	294	ProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIleHisGluValGln	313
Db	495	CCAGAGATGGGTGGCCAGGCTCACATGCCCTACACCACTGCCGTGATTCATGAGGTGCAG	436
Qy	314	ArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluVal	333
Db	435	CGCTTTGGGGACATCGTCCCTGGGTGTGACCCATATGACATCCCGTGACATCGAAGTA	376
Qy	334	GlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSerSerValLeuLys	353
Db	375	CAGGGCTCCCGATCCCTAAGGAAGCACATCATCCAAACCTGTCTCATCGGTCTGAAG	316
Qy	354	AspGluAlaValTrpGluLysProPheArgPheHisProGluHisPheLeuAspAlaGln	373
Db	315	GATGAGCCGCTCTGGGAGNAGCCCTTCGCTTCACCCCGNACACTTCCTGGATGCCCAG	256
Qy	374	GlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArgAlaCysLeu	393
Db	255	GGCCACTTTGTGAAGCCGAGGCCCTTCGTGCTTCTCAGCAGGCCCGCTGCATGCCTC	196
Qy	394	GlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeuLeuGlnHisPhe	413
Db	195	GGGAGGCCCTGGCCCGCATGAGCTTTCCTTCCTTCACCTCCCTGCTGCAGCACTTC	136
Qy	414	SerPheSerValProThrGlyGlnProArgProSerHisGlyValPheAlaPheLeu	433
Db	135	AGCTTCTCGGTGCCCACTGGACAGCCCCGGGCCAGCCACCATGTGTGTCTTGTCTTCCTG	76
Qy	434	ValThrProSerProTyrGluLeuCysAlaValProArg	446
Db	75	GTGACCCCATCCCTCATGAGCTTTGTGCTGTGCTGTGCCCCGC	37

RESULT 14	1088 bp	linear	EST 15-MAY-2003
BX432878			
LOCUS			
DEFINITION	BX432878 Homo sapiens	FETAL LIVER	Homo sapiens cDNA clone
	CSGD0004YJ13	5-PRIME.	MRNA sequence.

ORGANISM		Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
----------	--	--------------	--

REFERENCE 1 (bases 1 to 1088)

REFERENCES
I (Dases I CO 1988)
AUTHORS
Ij W.B., Gruber C., Jessee J. and Polaves D.

AUTHORS LI, W.B.; GRUBER, C.; JESSEE, J. and POLAYES, D.
TITLE Full-length cDNA libraries and normalization

TITLE	Full-length cDNA	Unpublished (2001)
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

JOURNAL Unpublished (2001)

COMMENT
Contact: Genoscope

Genoscope - Centre National de S

BP 191 91006 EVRY cedex - France

QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
 Db 667 CGCTTCTCAGGCTGCTGGACCTAGCTCAGGAGGACTGAAGGAGGAGTGGGCTTCTG 726
 QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisLeuProAlaLeuAlaGlyLysVal 189
 Db 727 CGGAGGTGCTGAATGCTGCTCCCGCTCTCCATATCCAGCGTGGCTGCAAGGTC 786
 QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
 Db 787 CTACGCTTCCAAAGGCTTTCSTGACCCAGCTGGATGAGCTGCTTAAGTGAACAGATG 846
 QY 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
 Db 847 ACCTGGACCCAGCCAGCCCGGAGACCTGACTGAGGCTTCTTGGCAGAGATGGAG 906
 QY 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgLysValValAla 249
 Db 907 AAGGCAAGGGGAACCTGAGAG---CGSTTCAATGATGAGAACTG-CGCATAGTGGTGGCT 962
 QY 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 269
 Db 963 GATCG-TTCTCTCCCGATGGTACCACTCG-ACCMSSTG-GACTGGGGGCTCTCGCTC 1019
 QY 270 MetIleLeuHisProAspValGlnArg 278
 Db 1020 ATGAT-CTACATCG-GATGTRCARCGC 1044

RESULT 15

BG743095

LOCUS

DEFINITION

602634230F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4779415 5',

mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 982)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-femail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
 Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLMI0636 row: 1 column: 08
 High quality sequence stop: 932.
 Location/Qualifiers

FEATURES

source

1..982
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4779415"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_Skn3"
 /note="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Alignment Scores:

Pred. No.:

Score:

4.32e-120 Length: 982
 1201.50 Matches: 255

Percent Similarity: 75.43% Conservative: 6
 Best Local Similarity: 73.70% Mismatches: 33
 Query Match: 51.57% Indels: 55
 DB: 12 Gaps: 3

US-09-820-788A-2 (1-446) x BG743095 (1-982)

QY 110 LeuGlyPheGlyProArgSer----- 116
 Db 42 TTGGGCTGGCAAGAGTCGTGGAGCAGTGGTGACCGAGGAGCGCGCTTGTGT 101
 QY 117 -----GlnGlyArgProPheArgProAsnGlyLeuLeuAspLysAla 130
 Db 102 GCGCCTTCCGCGACCAAGCGCGCCCTTTGCCCCAACGGCTCTTTGGACAAAGCC 161
 QY 131 ValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluThrPheAspProArg 150
 Db 162 GTGAGCAACGTATGCTCCCTCCCTCAGCTCGGGCGCGCTTCGAGTACGACCGCTCGC 221
 QY 151 PheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeuArg 170
 Db 222 TTCTCAGGCTGCTGACCTAGCTCAGAGGGAGCTGAAGAGGAGTTCGGCTTCTTCGCG 281
 QY 171 GluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysValLeu 190
 Db 282 GAGGTGCTGATGCTGCTCCCGCTCCAGCGCATCCAGCGCTGGCTGGCAAGTCTCTA 341
 QY 191 ArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrClnuHisArgMetThr 210
 Db 342 CGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGCTGCTTAACGAGCAGCATGAGGACC 401
 QY 211 TrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGluLys 230
 Db 402 TGGGACCCAGCCAGCCAGCCAGCTGACTGAGGCTTCTTGGCAAGAGGAGAG 461
 QY 231 AlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgLysValValAlaAsp 250
 Db 462 GCCAAGGGGAGCCCTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGT-GTC----- 514
 QY 251 LeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuMet 270
 Db 514 ----- 514
 QY 271 IleLeuHisProAspValGlnArgValGlnGluIleAspAspValIleGlyGln 290
 Db 515 -----CGTGTCCAAACAGGAGTGCACGACGCTGCACCATCCCGTGCATTCAC 550
 QY 291 ValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIleHis 310
 Db 551 GTGGCGCGACAGAGATGGTGCACGAGCTCGCATGCGCTGCACCATCCCGTGTATTCAC 610
 QY 311 GluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArgAsp 330
 Db 611 GAGGTGCAGCGCTTTGGGACATCATCCCTGAGTGTGACCCATATGACATCATCCCGTGAC 670
 QY 331 IleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSerSer 350
 Db 671 ATCGAAGTACAGGCTTCCGCATCCTTAAGGGACGACACTCATCAACCACTCTCATCG 730
 QY 351 ValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPheLeu 370
 Db 731 GTGTGAAGGATGAGCGCTGTGAAGAGCCCTTCCGCTTCCACCCCGCAACTTCTCTG 790
 QY 371 AspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArgArg 390
 Db 791 GGATGCCCAAGGACACTTTGTGAAGCGGAGGCTTCTTCTGCTTTCTCAGCGAGCGCGCT 850
 QY 391 AlaCysLeuGlyGluProLeuAlaArgMetGluLeuPhePhePheThrSerLeuLeu 410
 Db 851 GCATGCTCGGGAGCCCTGGCGCATGGGG-CTCTTTCTCTT-TTACCTCTCTTGTGC 908
 QY 411 GlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyValPhe 430

Db 909 -----AGCACTTCAGTTCTCGGTGGCGGGAAGGCCCGGCAGCACTCTCGTGTG 959
QY 431 AlaPheLeuValThrPro 436
Db :::||| |
960 TCAGTTTGGTGGACCCA 977

Search completed: February 29, 2004, 03:02:20
Job time : 2393 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 28, 2004, 16:42:05 ; Search time 395 Seconds

(without alignments)
4796.700 Million cell updates/sec

Title: US-09-820-788A-2

Perfect score: 2330

Sequence: 1 MGLEALVPLAIVAIALLLV.....HGVFAFLVTPSPYELCAVPR 446

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cn2 1/USPTO.spool/US09820788/runat_24022004.141407.9072/app.query.fasta_1.583
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human0.1
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09820788 @CGN 1 1 352 @runat_24022004.141407.9072 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Geneseq_29Jan04.*
- 1: Geneseq_1980s.*
- 2: Geneseq_1990s.*
- 3: Geneseq_2000s.*
- 4: Geneseq_2001as.*
- 5: Geneseq_2001bs.*
- 6: Geneseq_2002s.*
- 7: Geneseq_2003as.*
- 8: Geneseq_2003bs.*
- 9: Geneseq_2003cs.*
- 10: Geneseq_2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2330	100.0	1537	7 AAL53565	Aal53565 CDNA enco
2	2290.5	98.3	1567	7 ADB25779	Adb25779 Human CYP
3	2289.5	98.3	1494	2 AAO87730	Aao87730 Human aux
4	2289.5	98.3	1494	2 AAT28396	Aat28396 Human CYP
5	2288.5	98.2	1567	7 ADB25777	Adb25777 Human CYP
6	2286.5	98.1	1567	7 ACA61331	Aca61331 Human CYP
7	2286.5	98.1	1567	7 ADB25776	Adb25776 Human CYP
8	2285.5	98.1	1494	2 AAO87731	Aao87731 Human aux

9	2285.5	98.1	1494	2 AAT28397	Aat28397 Human CYP
10	2285.5	98.1	1494	2 AAT17388	Aat17388 Human der
11	2285.5	98.1	1494	6 ABQ72216	Abq72216 Human CYP
12	2285.5	98.1	1567	7 ADB25778	Adb25778 Human CYP
13	2282.5	98.0	1494	8 ACF06056	Acf06056 Human CYP
14	2282.5	98.0	1567	7 ACA61303	Aca61303 Human CYP
15	2281.5	97.9	1494	2 AAO87729	Aao87729 Human aux
16	2281.5	97.9	1494	2 AAT28395	Aat28395 Human CYP
17	2277.5	97.7	1494	2 AAO87732	Aao87732 Human aux
18	2277.5	97.7	1494	2 AAT28398	Aat28398 Human CYP
19	2271.5	97.5	1565	7 ACA61305	Aca61305 Human CYP
20	2267.5	97.3	1568	7 ACA61304	Aca61304 Human CYP
21	2250.5	96.6	1545	2 AAV19496	Aav19496 Cytochrom
22	2208	94.8	1586	2 AAO12893	Aao12893 Debrisoqu
23	2093.5	89.8	1593	4 AAD09937	Aad09937 Human dru
24	1708.5	73.3	2746	6 ABK63556	Abk63556 Rat seque
25	1708.5	73.3	2746	9 ADB53689	Adb53689 Primary r
26	1685	72.3	1449	6 ABN97193	Abn97193 Gene #369
27	1663	71.4	4500	7 ADB25775	Adb25775 Human CYP
28	1662	71.3	6472	6 ABQ72364	Abq72364 Human CYP
29	1662	71.3	9432	6 AAD34213	Aad34213 Human CYP
30	1662	71.3	9432	7 ACA61301	Aca61301 Human CYP
31	1651	70.9	9433	7 ACA61302	Aca61302 Human CYP
32	1641	70.4	1698	9 ADB58102	Adb58102 Toxicity
33	1638.5	70.3	1628	9 ADB56180	Adb56180 Toxicity
34	1638.5	70.3	1628	9 ADB50703	Adb50703 Primary r
35	1638.5	70.3	1628	9 ADB50703	Adb50703 Primary r
36	1615.5	69.3	1630	9 ADB59020	Adb59020 Toxicity
37	1615.5	69.3	1630	9 ADB53778	Adb53778 Primary r
38	1599	68.6	6472	6 ABQ72215	Abq72215 Human CYP
39	1551.5	66.6	10278	6 AAL53566	Aal53566 Genomic D
40	1525	65.5	1270	6 ABN97194	Abn97194 Gene #369
41	1141	49.0	2170	9 AAT33958	Aat33958 Human pig
42	1141	49.0	2170	9 ADC26596	Adc26596 Human CYP
43	1141	49.0	2170	9 ADC26591	Adc26591 Human CYP
44	1141	49.0	2170	9 ADC26793	Adc26793 Human lip
45	1132	48.6	2170	7 ABT33957	Abt33957 Human pig

ALIGNMENTS

RESULT 1

AAL53565
ID AAL53565 standard; cDNA; 1537 BP.

XX

AC AAL53565;

XX

DT 30-JAN-2003 (first entry)

XX

DE CDNA encoding the drug-metabolising enzyme protein.

XX

KW Cytostatic; antiparkinsonian; gene therapy; cancer; Parkinson's disease;

KW human drug-metabolising protein; enzyme; gene; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 78..1418

FT /*tag= a

FT /product= "Human drug-metabolising protein"

XX

PN WO200279233-A1.

XX

PD 10-OCT-2002.

XX

PF 01-APR-2002; 2002WO-US009738.

XX

PR 30-MAR-2001; 2001US-00820788.

XX

PA (PEKE) PE CORP NY.

PA (DFRA/) DI FRANCESCO V.

PA (BEAS/) BEASLEY E M.

XX

PI Shao W, Yan C;
 XX WPI; 2003-040649/03.
 DR P-PSDB; AAO26404.
 XX
 PT New human drug-metabolizing proteins and nucleic acids related to the
 PT Cytochrome P450 IID drug-metabolizing enzyme subfamily, useful for
 PT treating a condition mediated by a human enzyme protein e.g., cancer.
 XX
 XX Claim 4; Fig 1; 72pp; English.

XX The invention relates to a novel isolated polypeptide comprising a 446-
 CC amino acid sequence or its allelic variant, orthologue or fragment. The
 CC allelic variant or orthologue is encoded by a nucleic acid that
 CC hybridises under stringent conditions to the opposite strand of the
 CC nucleic acid comprising a sequence having 1537 or 10278 bp. The fragment
 CC comprises at least 10 contiguous amino acids of the 446-amino acid
 CC sequence. The polypeptide is useful for preparing a pharmaceutical
 CC composition for treating a disease or condition mediated by a human
 CC enzyme protein, e.g. cancer or Parkinson's disease. This polynucleotide
 CC sequence represents the cDNA encoding the human drug-metabolising protein
 CC of the invention

SQ Sequence 1537 BP; 301 A; 493 C; 443 G; 300 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.2e-228 Length: 1537
 Score: 2330.00 Matches: 446
 Percent Similarity: 100.00% Conservativeness: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0

US-09-820-788A-2 (1-446) x AAL53565 (1-1537)

QY 1 MetGlyLeuAlaLeuValProLeuAlaValAlaValAlaValAlaValAlaValAlaVal 20
 DB 78 ATGGGGCTAGAGCACTGTGGTCCCTGGCGGTGATAGTGCCCATCTTCTGCTCTCTGGTG 137
 QY 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeu 40
 DB 138 GACCTGATGACCGCGCGCAACGCTGGGCTGGACGCTACTACACGAGCCCTGCGACATG 197
 QY 41 ProGlyLeuGlyAsnLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
 DB 198 CCGGGCTGGGAACCTGCTGATGGACTTCCAGAACACACCATACCTGCTTCGACCCAG 257
 QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaValTrpProValValVal 80
 DB 258 TTGGGGCGCGCTTCGGGGAGCGTTCACGCTGCAGCTGGCTGGACGCGGTGGTGGTG 317
 QY 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
 DB 318 CTCAATGGGCTGGCGCGCGTGGCGAGGGCTGGTGACCCACCGCGAGGACACCGCGAC 377
 QY 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGlnGlyArgPro 120
 DB 378 CGCCCGCCCTGTCCCATCACCGATCCTGGGTTTTTGGGCGCGCTTCCCAAGGACGCCCC 437
 QY 121 PheArgProAsnGlyLeuLeuAspLysAlaValSerAsnValIleAlaSerLeuThrCys 140
 DB 438 TTTGCGCCCAACGGTCTCTTGACAAAGCCGTGAGCAACGTATCGCTCCCTCCCTGAC 497
 QY 141 GlyArgArgPheGlyTyrAspAspProArgPheLeuArgLeuAspLeuAlaGlnGlu 160
 DB 498 GGGCGCGCGCTTCAGGTACACGACACCGCTCTCTCTCAGGCTGCTGAGACCTAGCTCAGGAG 557
 QY 161 GlyLeuLysGluGluSerGlyPheLeuArgGluValLeuAsnAlaValProValLeuLeu 180
 DB 558 GGACTGAAGAGGAGTTCGGGCTTTCTGGCGGAGGTGCTGAATGCTGTCCCGCTCTCTG 617
 QY 181 HisIleProAlaLeuAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeu 200

DB 618 CATATCCAGCGCTGGCGCAAGGTCCTACGCTTCCAAAGGCTTTCTCTGACCCAGCTG 677
 QY 201 AspGluLeuLeuThrGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeu 220
 DB 678 GATGAGCTGCTAACTGAGCACAGGATGACCTGGGAGCCAGCCGCCCGGAGACCTG 737
 QY 221 ThrGluAlaPheLeuAlaGluMetGluLysAlaLysGlyAsnProGluSerSerPheAsn 240
 DB 738 ACTGAGGCTTCTCTGGCAGAGATGGAGAAGCCCAAGGGGAAACCTTGAGAGAGCTTCAAT 797
 QY 241 AspGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetValThrThrSer 260
 DB 798 GATGAGAACCCTGGCATAGTGGTGGCTGACCTGTTCTCTGCGGGATGCTGACCACTCG 857
 QY 261 ThrThrLeuAlaTrpGlyLeuLeuMetIleLeuHisProAspValGlnArgArgVal 280
 DB 858 ACCAGCTGGGCTGGGCGCTCTGCTCATGATCCTACATCGGATGTGTCAGCGCGCTGC 917
 QY 281 GlnGlnGluIleAspAspValIleGlyGlnValArgArgProGluMetGlyAspGlnAla 300
 DB 918 CAACAGGAGATCGACGACGTATAGGCGAGGTGGCGGACGAGATGGGTGACCGAGCT 977
 QY 301 HisMetProTyrThrThrAlaValIleHisGluValGlnArgPheGlyAspIleValPro 320
 DB 978 CACATGCCCTACACCACTGCCGTGATTCTATCAGGTGCGAGCGCTTTGGGGACATCGTCCC 1037
 QY 321 LeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPheArgIleProLys 340
 DB 1038 CTGGGTGTGAGCCCATATGACATCCGTCGACATCGAAGTACAGGGCTTCCGCATCCCTAAG 1097
 QY 341 GlyThrThrIleThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLys 360
 DB 1098 GGAACGACATCATCACCACTGTCTCATCGGTGTGAAGGATGAGGCGCTCTGGGAGAG 1157
 QY 361 ProPheArgPheHisProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGlu 380
 DB 1158 CCTTCCGCTTCCACCCCGACACTTCTGATGCCAGGCGCCACTTTGTGAAGCCGAG 1217
 QY 381 AlaPheLeuProPheSerAlaGlyArgArgAlaCysLeuGlyGluProLeuAlaArgMet 400
 DB 1218 GCCTTCTCTGCTTTCTCAGCAGGCGCGCTGCATGCTCGGCGGAGCCCTGCGCGCATG 1277
 QY 401 GluLeuPheLeuPheThrSerLeuLeuGlnHisPheSerPheSerValProThrGly 420
 DB 1278 GAGCTTCTCTCTTCTTCCCTCCCTGCTGTCAGGACCTTACGCTTCTGCTGCCACTGGA 1337
 QY 421 GlnProArgProSerHisHisGlyValPheAlaPheLeuValThrProSerProTyrGlu 440
 DB 1338 CAGCCCGCGCCAGCCACCATGCTGTTCTTCTGCTTCTGCTGACCCCATCCCTATGAG 1397
 QY 441 LeuCysAlaValProArg 446
 DB 1398 CTTTGTGCTGCGCCCGC 1415
 RESULT 2
 ADB25779
 ID ADB25779 standard; DNA; 1567 BP.
 XX
 AC ADB25779;
 DT 20-NOV-2003 (first entry)
 DE Human CYP2D6-related gene #4.
 KW human; mutant CYP2D6 gene; drug analysis; drug testing; gene; ds.
 OS Homo sapiens.
 XX
 FH Key
 FT CDS
 FT CDS
 FT CDS
 XX Location/Qualifiers
 1..1494
 /*tag= a
 /product= "Human CYP2D6-related protein #4"

QY 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
DB 1201 TCGTGCTGCTGAAGATGAGCGCTCTGGAGAGACCTTCGCTTCCACCCGGAACACTTC 1260
QY 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
DB 1261 CTGATGCCAGGCGGCACCTTTGTGAAGCCGAGAGGCTTCTGCTTTCTCAGAGGCCGC 1320
QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrSerLeu 409
DB 1321 CGTGATGCCCTGGGAGGCCCTTGGCCGCGCATGGAGCTCTTCTCTTCTCACCCTCCCTG 1380
QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
DB 1381 CTGCACACTTCAGCTTCTGGTGCCCACTGGAGACGCCGCGCCAGCCACCATGGTGTC 1440
QY 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
DB 1441 TTTGCTTTCTGGTGACCCCATCCCTATGAGCTTTGTGCTGTGCGCCGC 1491

RESULT 4

AAT28396
ID AAT28396 standard; DNA; 1494 BP.
XX AAT28396;
XX
DT 11-OCT-1996 (first entry)
XX Human cytochrome P450 molecular species 2D6 variant #2 gene.
DE Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;
KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;
KW evaluation; safety; fusion protein; metabolic; detoxification;
KW carcinogenic; ds.
XX Homo sapiens.
XX JP08056695-A.
XX 05-MAR-1996.
XX 15-JUL-1994; 94JP-00164184.
XX 20-JUL-1993; 93JP-00201120.
PR 30-JUL-1993; 93JP-00208279.
PR 17-JUN-1994; 94JP-00136053.
XX (SUMO) SUMITOMO CHEM CO LTD.
XX WPI; 1996-182311/19.
DR P-PSDB; AAR93183.
XX
PT Novel method for the evaluation of the safety of a cpd. - using a human
PT cytochrome P450 and yeast NADPH reductase to determine whether the
PT analyte cpd. is detoxified or metabolised to a carcinogen.
PS Example 1; Page 51-53; 74pp; Japanese.
XX
CC This is the nucleotide sequence of the human cytochrome P450 molecular
CC species 2D6 variant #2 which encodes a protein of 497 amino acids. The
CC gene was amplified from a human liver derived cDNA library as 2 fragments
CC of 0.4 and 0.9 kb using primers AAT26953-6. The prod. was cloned into the
CC yeast expression vector pAAH5N to generate plasmid p2D6 for prodn. of the
CC cytochrome only or into the vector pAHR to generate the plasmid p2C6r
CC for co-prodn. with the yeast NADPH-P450 reductase. The sequence is placed
CC under control of the yeast ADH gene promoter and terminator. The vectors
CC are used in a method for evaluating the safety of a cpd. by reacting the
CC test cpd. with recombinantly produced human cytochrome P450 mol. species
CC 1A2 (AAT28380), 2C9 (AAT28381), 2E1 (AAT28382), 3A4 (AAT28383) or their
CC variants (AAT28384-98) together with yeast NADPH-P450 reductase (either
CC as a fused protein or as a cell extract) and analysing the resultant
CC metabolite. The cpd. is considered "safe" if it is detoxified or not

CC rendered carcinogenic or "unsafe" if it is not detoxified or is
CC metabolised to a carcinogenic cpd

XX
SQ Sequence 1494 BP; 248 A; 509 C; 446 G; 291 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,91e-224 Length: 1494
Score: 2289.50 Matches: 445
Percent Similarity: 89.54% Conservatives: 0
Best Local Similarity: 89.54% Mismatches: 1
Query Match: 98.26% Indels: 51
DB: 2 Gaps: 1

US-09-820-788A-2 (1-446) x AAT28396 (1-1494)

QY 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
DB 1 ATGGGGCTAGAAAGCACTGGTGCCCTGGCGCTGATAGTAGGCCCATCTTCTGCTCTCTGGTG 60
QY 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40
DB 61 GACCTGATGCACGGCGCCCAACGCTGGGCTGCACGCTACCCACGAGGCCCTGCCACTG 120
QY 41 ProGlyLeuGlyAsnLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
DB 121 CCGGGCTGGGCAACCTGCTGCATGTGGACTTCCAGAACACACCATACTGCTTCGACCAG 180
QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
DB 181 TTGGGGCGCCGCTTCGGGACGCTGTCAGCTTGGAGCTGGCGCGGGTGGTGGTGGTGGTGG 240
QY 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
DB 241 CTCATGGCTGGCGGCGCTGGCGAGGGCTGGTGACCCACGCGGAGGACACCGCGGAC 300
QY 101 ArgProValProLeuThrGlnIleLeuGlyPheGlyProArgSerGln 117
DB 301 CGCCGCGCTGTGCCCATCACCAGATCCTGGGCTTTCGGGCGCGCTTCCCAAGGGGTGTTTC 360
QY 117 ----- 117
DB 361 CTGGCGCGCTATGGGCGCGCTGGGCGGAGAGAGCGCTTCTCGCTTCCACCTTGGCGC 420
QY 117 ----- 117
DB 421 AACTTGGGCTGGGCAAGAGTCGCTGGAGAGTGGGTGACCGAGAGGAGCGCGCTTCCCTT 480
QY 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
DB 481 TGTGCGCGCTTCGCCAACCACTCCGAGCGCCCTTTCGCCCAACGGTCTCTTGGACAAA 540
QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluThrAspAspPro 149
DB 541 CGCGTGAGCAAGCTGATCGCTTCCCTGCGGGCGCGCTTCCGAAATACGACGACCTT 600
QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLeuGluGluSerGlyPheLeu 169
DB 601 CGCTTCTCAGGCTGCTGGACCTAGCTCAGGAGGAGACTGAAGAGGAGTCTGGGCTTCTG 660
QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
DB 661 CGCAGGAGTGTGAATGCTGTCCCGCTCTCTGTCATATCCACGCGCTGGCTGGCAAGGTC 720
QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrCysArgMet 209
DB 721 CTACCTTCCAAAAGGCTTTCCTGACCCAGCTGGATGAGCTGCTAACTGAGCAGAGATG 780
QY 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
DB 781 ACCTGGGACCCAGCCCGCCCGCCCGACCTGACTGAGGCGCTTCTGGCAGAGATGGAG 840
QY 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249

```
Db 841 AAGGCCAAGGGGAACCCCTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGGTCT 900
Qy 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeu 269
Db 901 GACCTGTTCTCGCCGGATGTGACCACTCGACACGCTGGCTGGGGCTCTCTGCTC 960
Qy 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGlnIleAspValIleGly 289
Db 961 ATGATCTTACATCCGATGTGAGCGCGCTGTCCAAACAGGAGATCGACGCTGATAGG 1020
Qy 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyThrThrAlaValle 309
Db 1021 CAGGTGCGCGCAGACAGATGGGTGACCAAGGTTCATGCCCTCACCACTCCCGTGAT 1080
Qy 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATGAGGTGCACGCTTTGGGAGATCGTCCCTCGGGTGTGACCATATGACATCCCGT 1140
Qy 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAAGTACAGGGCTTCGCGATCCCTAAGGGAACGACACTCATCACCAACCTGTCA 1200
Qy 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAGGATGAGCGCTCTGGGAGAGCCCTTCGCTTCCACCCCGAACACTTC 1260
Qy 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGGCCACTTTGTGAAGCCGGAGGCTTCCTGCCCTTCTCAGCAGGCGGC 1320
Qy 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeu 409
Db 1321 CGTGATGCTCGGGAGCCCTTCGCGCATGGAGCTCTTCCTCTTCTTCACTCCCTG 1380
Qy 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
Db 1381 CTGACGACATTTCAGCTTCTCGGTGCCCACTGGACAGCCCGGCCAGCCACCATGGTCTC 1440
Qy 430 PheAlaPheLeuValThrProSerProTyThrGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTCTCGGTGACCCCACTCCCTCTATGAGCTTTGTGTGTGCCCGGC 1491
```

RESULT 5

ADB25777

ID ADB25777 standard; DNA; 1567 BP.

XX ADB25777;

AC ADB25777;

XX 20-NOV-2003 (first entry)

DT Human CYP2D6-related gene #2.

XX human; mutant CYP2D6 gene; drug analysis; drug testing; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1494

XX /*tag= a

XX /product= "Human CYP2D6-related protein #2"

XX WO2003050282-A1.

XX 19-JUN-2003.

XX 05-DEC-2002; 2002WO-JP012748.

XX 06-DEC-2001; 2001JP-00372548.

XX (TSUR) TSUMURA & CO.

XX Taniyama M, Ogawa K, Tsuchiya N, Hibino T;

XX

```
DR WPI; 2003-505401/47.
DR P-PSDB; ADB25832.
XX Genetic polymorphisms of CYP2D6 gene in human population for analysis of
PT drug effect on individual patients and testing of new drugs.
XX Claim 8; Page 43-46; 75pp; Japanese.
XX The invention comprises mutant forms of the human CYP2D6 gene, containing
CC one or more of the following mutations G125A, C1858T, T2874C and C2875T.
CC The mutant human CYP2D6 genes of the invention are useful for analysing
CC the effect of drugs on individual patients and testing of new drugs. The
CC present DNA sequence represents a human gene of the invention.
XX SQ Sequence 1567 BP; 266 A; 527 C; 464 G; 310 T; 0 U; 0 Other;
```

Alignment Scores:

```
Pred. No.: 9,36e-224 Length: 1567
Score: 2288.50 Matches: 445
Percent Similarity: 89.54% Conservative: 0
Best Local Similarity: 89.54% Mismatches: 1
Query Match: 98.22% Indels: 51
DB: 7 Gaps: 1
```

US-09-820-788A-2 (1-446) x ADB25777 (1-1567)

```
Qy 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
Db 1 ATGGGGCTAGAAAGCACTGGTCCCTGGCCGTGTAGTGGCCATCTTCTGCTCTCTGTTG 60
Qy 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuLeu 40
Db 61 GACCTGATGACACGGCGCAACGCTGGGCTGCACGCTACTCACACGAGGCCCTTGCACATG 120
Qy 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyThrCysPheAspGln 60
Db 121 CCGGGCTGGGCAACCTGCTGATGTGACTTCAGAAACACACACCATACTGCTTCGACCCAG 180
Qy 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValVal 80
Db 181 TTGGGGCCCGCTTCGGGGAGCGTTCAGCTGACGCTGGAGCTGGACCGCGTGGTCTGTTG 240
Qy 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTCAATGGCTGGCGGCGCTGGCGAGGCGTGGTGACCCACGCGAGAGACACCCGCCGAC 300
Qy 101 ArgProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln----- 117
Db 301 CGCCCGCTGTGCCCATCACCCAGATCCTGGGTTTGGGGCGCGTTCCCAAGGGGTGTTTC 360
Qy 117 ----- 117
Db 361 CTGGCGCGCTATGGGCGCGCTGGCGGAGCAGAGCGCTTCTCGTCTCCACCTTGGCGC 420
Qy 117 ----- 117
Db 421 AACTTGGCCCTGGGCAAGAGTCCGCTGGAGCAGTGGGTGACCGAGGAGGCCCTTGCCTT 480
Qy 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
Db 481 TGTGGCGCTTCGCCAACCACTCCGAGCGCCCTTTTCGCCCAACCGTCTCTTGGACAAA 540
Qy 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyThrAspAspPro 149
Db 541 GCCGTGAGCAACGTGATCGCTCCCTCACCTGCGGGCGCGCTTCGAGTACGACACCTT 600
Qy 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
Db 601 CGCTTCTCAGGCTGTGGACCTAGCTCAGAGAGGAGCTGAGAGGAGAGTCTGGGCTTTCTG 660
Qy 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
Db 661 CGCAGGGTGTGAATGCTGTCCCGTCTCTCTGATATATCCAGCGCTGGCTGGCAAGGTTC 720
```

QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuThrGluHisArgMet 209
DB 721 CTACGCTTCCAAAGGCTTCTCTGACCGAGCTGGATGCTAACTGAGCAGCATG 780
QY 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuGluMetGlu 229
DB 781 ACTGGGAGCCAGCCAGCCCGCCGAGACCTGACTGAGGCTTCTGGCAGAGATGGAG 840
QY 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
DB 841 AAGCCCAAGGGAACTCTGAGACAGCTTCAATGATGAGAACCTGGCGATAGTGGTGGCT 900
QY 250 AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTrpGlyLeuLeuLeu 269
DB 901 GACCTGCTTCTGCGGGATGGTGACCACTGACCACTGGCGCTGGCGCTCTCTGCTC 960
QY 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspValIleGly 289
DB 961 ATGATCTCATCTCCGATGTGACGCGCGGTGTCACAGGAGATCGACGCTGATGAGG 1020
QY 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
DB 1021 CAGGTGCGGCGACAGATGGTGACCACTGAGGCTCATGCTTACCACTGCGGTGAT 1080
QY 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
DB 1081 CATAGGTGACGCGCTTGGGACATCGTCCCGCTGGGTGTGACCCATATGACATCCCT 1140
QY 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrLeuIleThrAsnLeuSer 349
DB 1141 GACATCGAAGTACAGGGCTTCGCGCATCCCTAAGGGAACGACACTCATCAACCACTGTCA 1200
QY 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
DB 1201 TGGTGCTGAGATGAGCGCTTCTGGAGAGCCCTTCCGCTTCCACCCGACACTTC 1260
QY 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
DB 1261 CTGGATGCCAGGGCCACTTTGTGAAGCGGAGGCTTCTGCTTTCTCAGCAGCGCGC 1320
QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 409
DB 1321 CCGTATGCTCTCGGAGCGCCCTGGCGCGCATGGAGCTTCTCTTCTACCTCCCTG 1380
QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
DB 1381 CTGACGACTTCAGCTTCTCGGTGCCCTGACAGACGCGCGCCGACCATGCTGTC 1440
QY 430 PheAlaPheLeuValThrProSerProTyrGluLeuLysCysAlaValProArg 446
DB 1441 TTTGCTTTCTGCTGACCCCATCCCTATGAGCTTTGCTGTGCGCGCGC 1491

RESULT 6
ID ACA61331
XX AC ACA61331 standard; cDNA; 1567 BP.
XX AC ACA61331;
XX 16-JUL-2003 (first entry)
XX Human cytochrome p450 gene CYP2D6 G5799C variant, cDNA.
KW Human; ss; gene: cytochrome P450; CYP2D6; chromosome 22; SNP;
KW single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
XX psychiatric disorder; drug sensitivity.
OS Homo sapiens.
FH Key Location/Qualifiers
FT 1..1494
FT /*tag= a
FT /product= "CYP2D6 protein"

variation replace(1457,G)
/*tag= b
/standard_name= "Single nucleotide polymorphism"
EP1281755-A2.
05-FEB-2003.
16-JUL-2002; 2002EP-00254972.
31-JUL-2001; 2001US-0309111P.
(PFIZ) PFIZER PROD INC.
Milos PM; Webb SM;
WPI: 2003-373769/36.
P-PSDB; AB09598.
New cytochrome P450 2D6 gene variants and polypeptides, useful for
determining if a subject has or is at risk of developing a drug
sensitivity condition or disorder that is associated with an aberrant
CYP2D6 activity.

Disclosure; Fig 7; 89pp; English.

The invention relates to an isolated nucleic acid comprising a cytochrome
P450 2D6 gene variant, e.g. G5799C or C5816AT (referring to the genomic
sequence or the same variant nucleotide in the corresponding cDNA
sequences). Also included are probes, primers (allele specific
oligonucleotides) and arrays used to detect and/or amplify the CYP2D6
gene polymorphic regions, the variant polypeptides, antibodies which are
capable of distinguishing between the variant and wild-type polypeptides,
determining whether a subject has a genetic deficiency for metabolising a
drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
determining whether an individual is susceptible to being a poor
metaboliser of drugs. The DNA probe is useful for hybridising to a
variant form of the CYP2D6 gene. The primer is useful for amplifying the
C5816TA allelic variant. The allele specific nucleotide is useful for the
detection of the C5816TA allelic variant. The methods are useful for
determining whether a subject has a genetic deficiency for metabolising a
drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and
determining if an individual is susceptible to being a poor metaboliser
of drugs. The nucleic acids are useful as probes or primers for
determining whether a subject has a genetic deficiency for metabolising
drugs that are substrates of P450 CYP2D6. The methods are useful for
determining if a subject has or is at risk of developing a drug
sensitivity condition or disorder that is associated with an aberrant
CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
appropriate drugs or determining the course of treatment to administer to
a subject to treat cardiovascular or psychiatric disorders, or for
treating a subject with a drug sensitivity or disorder associated with a
specific allelic variant of a polymorphic region of the CYP2D6 gene. The
antibodies are useful for monitoring CYP2D6 protein levels in an
individual for determining whether a subject has a disease or conditions
associated with an aberrant CYP2D6 protein level. The gene is located on
human chromosome 22. The present sequence is the CYP2D6 cDNA carrying the
G5799C/C5816TA double variation

Sequence 1567 BP; 267 A; 528 C; 464 G; 308 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.5e-223 Length: 1567
Score: 2286.50 Matches: 444
Percent Similarity: 89.54% Conservat: 1
Best Local Similarity: 89.34% Mismatches: 1
Query Match: 98.13% Indels: 51
DB: 7 Gaps: 1

US-09-820-788A-2 (1-446) x ACA61331 (1-1567)

QY 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20

```
Db 1 ATGGGGCTAGAGACACTGGTGGCCCTGGCGGTGATAGTGCCCATCTTCTCGTCTCTGGTG 60
Qy 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeu 40
Db 61 GACCTGATGACACGGCGCCCAACGCTGGGTGACGCTACCCACGAGGCCCTGCGCACTG 120
Qy 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 121 CCCGGCTGGGCAACTGCTGATGTGGACTTCCAGAACACACCATCTGCTTCGACCAG 180
Qy 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
Db 181 TTGCGGCGCGCTTCCGGGAGCGTGTTCAGCTGCGAGCTGGCTGACGCCGCTGGTGGTG 240
Qy 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTCAATGGGCTGGCGCGCTGGCGAGCGCTGGTGACCCACGCGGAGACACACCGCCGAC 300
Qy 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln----- 117
Db 301 CGCCGCGCTGTGCCCATCACCCAGATCCTGGGTTCGGGCGCGCTTCCAGGGGTGTTC 360
Qy 117 ----- 117
Db 361 CTGGCGCGCTATGGGCGCGCTGGCGGAGAGAGCGGCTTCTCGTGTCCACCTTGGCG 420
Qy 117 ----- 117
Db 421 AACTTTGGGCTGGGCAAGAGTCTGGGAGCAGTGGGTGACCGAGGCGCGCTTGCCTT 480
Qy 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
Db 481 TGTGGCGCTTGGCAACCACTCCGAGCGCCCTTTTCGCCCAACCGGTCTTCTTGACAAA 540
Qy 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGlyTrpAspPro 149
Db 541 GCGGTGAGCAACGTATCGCTCCCTCACCTGGCGGCGCGCTTCGAGTACGACGACCT 600
Qy 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
Db 601 CGCTTCTCCTCAGGCTGTGGACCTAGCTCAGGAGGAGCTCAAGAGGAGGTCTGGGCTTCTG 660
Qy 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyVal 189
Db 661 CCGAGGAGTCTGAATGCTGTCCCGTCTCTGACATATCCACGCGCTGGCTGGCAAGGTC 720
Qy 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
Db 721 CTACGCTTCCAAAGAGCTTCTGACCCAGCTGGATGAGCTGCTAACTGAGCAGCAGGATG 780
Qy 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 781 ACCTGGGACCCAGCCAGCCCGCCGAGACCTGACTGAGGCTTCTTGGCAGAGATGGAG 840
Qy 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
Db 841 AAGGCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACTCGCATAGTGGTGGCT 900
Qy 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 269
Db 901 GACCTGTCTCTGCGGGGATGTGACCACTGACCACTGCGTGGCTTGGGCTCTCTGCTC 960
Qy 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspAspValIleGly 289
Db 961 ATGATCTTACATCCGAGTGTGAGCGCGCTGTGTCACACAGGAGATCGACGAGTGTAGGG 1020
Qy 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
Db 1021 CAGGTGGCGCGACAGAGATGGGTGACCAAGGCTCACATGCGCCCTACACCACTGCGGTATT 1080
Qy 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
```

```
Db 1081 CATGAGGTGACAGCGCTTTGGGGACATCGTCCCGCTGGGTATGACCCATATGACATCCCGT 1140
Qy 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAAGTACAGGGCTTCCGCATCCCTAAGGGAGACGACTCATCAACACCTGTCA 1200
Qy 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAAGATGAGCGCGTCTGGGAGAGCGCTTCCGCTTCCACCCGACACTTC 1260
Qy 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGCGCACTTTGTGAAGCGGAGGCGCTTCTCGCTTCTCAGCAGCGCGC 1320
Qy 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrSerLeu 409
Db 1321 CGTGATCGCTCGGGGAGCGCTGCGCGCATGGAGCTTCTTCTTCTTCCCTTCCCTCCCTG 1380
Qy 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
Db 1381 CTGACGCACTTTCAGCTTCTCGGTGCCACTGGACAGCCCGCGCCAGCCACCATGGTGTTC 1440
Qy 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTCTCTGTGACCCCATCCCTATGAGCTTTGTGTGTGCTGTGCCCCCGC 1491
RESULT 7
ADB25776
ID ADB25776 standard; DNA; 1567 BP.
XX
AC ADB25776;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human CYP2D6-related gene #1.
XX
KW human; mutant CYP2D6 gene; drug analysis; drug testing; gene; ds.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FH CDS 1..1494
FT /*tag= a
FT /product= "Human CYP2D6-related protein #1"
XX
PN WO2003050282-A1.
XX
PD 19-JUN-2003.
XX
PF 05-DEC-2002; 2002WO-JP012748.
XX
PR 06-DEC-2001; 2001JP-00372548.
XX
PA (TSUR ) TSUMURA & CO.
XX
PI Taniyama M, Ogawa K, Tsuchiya N, Hibino T;
XX
WPI; 2003-505401/47.
XX
DR P-PSDB; ADB25831.
XX
PT Genetic polymorphisms of CYP2D6 gene in human population for analysis of
PT drug effect on individual patients and testing of new drugs.
XX
PS Claim 8; Page 39-42; 75pp; Japanese.
XX
CC The invention comprises mutant forms of the human CYP2D6 gene, containing
CC one or more of the following mutations G125A, C1858T, T2874C and C2875T.
CC The mutant human CYP2D6 genes of the invention are useful for analysing
CC the effect of drugs on individual patients and testing of new drugs. The
XX present DNA sequence represents a human gene of the invention.
XX
SQ Sequence 1567 BP; 266 A; 527 C; 464 G; 310 T; 0 U; 0 Other;
```


XX (SUMO) SUMITOMO CHEM CO LTD.
XX PA Hayashi K, Sakaki T, Yabusaki Y, Komai K, Kaneko H, Nakatsuka I;
XX PI WPI; 1995-116991/16.
XX DR P-PSDB; AAR72377.
XX PS
XX PT Evaluation of safety of a chemical cpd. - using recombinant yeast
XX PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase.
XX PS
XX PS Example; Page 91-93; 124pp; English.
XX CC The nucleotide sequence of the cDNA coding region for the human auxiliary
XX CC cytochrome P450 species 2D6 variant 2. The gene contains variations at
XX CC bases 886: T to C and 1457: C to G as compared to the wild type sequence
XX CC (AAQ87729). These change the amino acid residues 296: Cys to Arg and 486:
XX CC Thr to Ser. The cDNA was amplified by PCR using the primers AAQ87763-6.
XX CC The product was cloned into the yeast expression vectors pAAH5N or pAHR
XX CC to produce the vectors p2D6 variant 2 for the expression of the
XX CC cytochrome P450 alone or p2D6 variant 2 for co-expression with the yeast
XX CC NADPH-P450 reductase. The vectors are used in a method for evaluating the
XX CC safety of a chemical compound by reacting the chemical compound with
XX CC recombinantly produced human cytochrome P450 molecular species 1A2
XX CC (AAQ87714), 2C9 (AAQ87715), 2E1 (AAQ87716), or 3A4 (AAQ87717) or their
XX CC auxiliary species and variants (AAQ87718-32), and yeast NADPH-P450
XX CC reductase, either as a fused protein or in cell extracts, and analysing
XX CC the resulting metabolite to assess the safety of the chemical compound.
XX CC The method is useful for determining whether the chemical compound, or
XX CC its metabolite, will be converted into a carcinogenic or mutagenic form
XX CC through metabolism in the liver. (Updated on 25-MAR-2003 to correct PN
XX CC field.)
XX SQ Sequence 1494 BP; 248 A; 508 C; 447 G; 291 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,77e-223 Length: 1494
Score: 2285.50 Matches: 444
Percent Similarity: 89.54% Conservative: 1
Best Local Similarity: 89.34% Mismatches: 1
Query Match: 98.09% Indels: 51
DB: 2 Gaps: 1
US-09-820-788a-2 (1-446) x AAQ87731 (1-1494)
QY 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
DB 1 ATGGGGCTAGAACGACCTGTCGCCCTGGCCGTGATAGTGGCCATCTTCTGCTCTGCTG 60
QY 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeu 40
DB 61 GACCTGATGCACCGCGCAACGCTGGGCTGCACGCTACCCACCGCCCTGCACTG 120
QY 41 ProGlyLeuGlyAsnLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
DB 121 CCCGGGCTGGGCAACCTGCTGATGTGACTTCCAGACACACCACTATGCTTCGACCAG 180
QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValVal 80
DB 181 TTGCGGCGCGCTTCGGGACGCTGTTACGCTGCAGCTGGCTGGCGCCGCTGGTGGT 240
QY 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
DB 241 CTCATATGGGCTGGCGCGCTGGCGAGGCGCTGGTGACCCACCGCGGAGACACCGCCGAC 300
QY 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln----- 117
DB 301 CGCCCGCTGTGCCCATCACCGATCTTGGGTTTCGGGCGCGCTTCCAGGGGTGTTTC 360
QY 117 ----- 117
DB 361 CTGGCGCGCTATGGGCGCGCTGGCGGAGAGAGCGCTTCTCCGCTCTCCACCTTGGCG 420

QY 117 ----- 117
DB 421 AACTTGGGCTGGGCAAGAGTGCCTGGAGCAGTGGGTGACGAGGAGCGCCCTGCTT 480
QY 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
DB 481 TGTGCGGCTTTCGCCAACCACTCCGAGCGCCCTTTCGCCCAACCGTCTCTTGGACAAA 540
QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149
DB 541 GCGGTGAGCAACGATGCTCCCTCACCTGCGGCGCGCTTTCGATATACGACGACCT 600
QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluSerGlyPheLeu 169
DB 601 CGCTTCTTCAGCTGCTGAGACTAGCTCAGAGGAGCTGAAGAGGAGTTCGGCTTCTG 660
QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
DB 661 CGCGAGGTGCTGAATGCTGCTCCCGTCTCTCGATATCCCGAGCGCTGGTGGCAAGTCT 720
QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
DB 721 CTACGCTTCCAAAAGGCTTTCCTGACCCAGCTGGATGAGTCTAACTGAGCAGCAGGATG 780
QY 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
DB 781 ACCTGGGACCGAGCCCGAGCTGCTGAGGCTGCTGAGGCTTCTCTGGCAGAGATGGAG 840
QY 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
DB 841 AAGGCCAAGGGAACCTCGAGCAGCTTCAATGATGAGAACCTCGCGCATAGTGGTGGCT 900
QY 250 AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTrpGlyLeuLeu 269
DB 901 GACCTGTCTCTGCGGGATGGTGACCACTCGACCGCTGGGCTGGGCGCTCTCTGCTC 960
QY 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspAspValIleGly 289
DB 961 ATGATCTTACATCCGATGTCAGCGCGCTGTCACAGGAGATCGACGAGTGTATAGGG 1020
QY 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
DB 1021 CAGGTGGCGGACAGAGATGGGTGACCGAGCTCACATGCTCACCACTGCGGTGAT 1080
QY 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
DB 1081 CATGAGGTGACGCGCTTGGGGACATCGTCCCTCGGTGTGACCCCATATGACATCCCGT 1140
QY 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
DB 1141 GACATCGAAGTACAGGCTTCCGATCCCTAAGGGAACGACACTCATCACCACCTGTCA 1200
QY 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
DB 1201 TCGGTGCTGAAGGATAGGCGCTCTGGGAGAGCCCTTCGCTTCCACCCCGAACACTTC 1260
QY 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
DB 1261 CTGGATGCCAGGCGCACTTTGTGAAGCGGAGGCTTCTCTGCTTCTCAGCAGCGCGC 1320
QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrSerLeu 409
DB 1321 CGTGATGCTTCGGGAGCGCCCTGGCCCGCATGAGCTTCTTCTTCTTCTTCTTCTTCT 1380
QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
DB 1381 CTGCAGCACTTTCAGCTTCTCGGTGCCACTGGACAGCGCGCGCGCGCAGCCACCATGTGTC 1440
QY 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
DB 1441 TTTGCTTTCTCTGGTGAGCCCATTCCTCCCTATGAGCTTTGTGTGTGCTGCCCCGC 1491
RESULT 9

AAT28397
 ID AAT28397 standard; DNR; 1494 BP.
 XX
 AC AAT28397;
 XX
 DT 11-OCT-1996 (first entry)
 XX
 DE Human cytochrome P450 molecular species 2D6 variant #3 gene.
 XX
 KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;
 KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;
 KW evaluation; safety; fusion protein; metabolite; detoxification;
 KW carcinogenic; ds.
 XX
 OS Homo sapiens.
 XX
 PN JP08056695-A.
 XX
 PD 05-MAR-1996.
 XX
 PF 15-JUL-1994; 94JP-00164184.
 XX
 PR 20-JUL-1993; 93JP-00201120.
 PR 30-JUL-1993; 93JP-00208279.
 PR 17-JUN-1994; 94JP-00136053.
 XX
 PA (SUMO) SUMITOMO CHEM CO LTD.
 XX
 DR WPI; 1996-182311/19.
 DR P-PSDB; AAR93184.
 XX
 PT Novel method for the evaluation of the safety of a cpd. - using a human
 PT cytochrome P450 and yeast NADPH reductase to determine whether the
 PT analyte cpd. is detoxified or metabolised to a carcinogen.
 XX
 PS Example 1; Page 53-55; 74pp; Japanese.
 XX
 CC This is the nucleotide sequence of the human cytochrome P450 molecular
 CC species 2D6 variant #3 which encodes a protein of 497 amino acids. The
 CC gene was amplified from a human liver derived cDNA library as 2 fragments
 CC of 0.4 and 0.9 kb using primers AAT26953-6. The prod. was cloned into the
 CC yeast expression vector pAAH5N to generate plasmid p2D6 for prodn. of the
 CC cytochrome only or into the vector pAHRH to generate the plasmid p2C6r
 CC for co-prodn. with the yeast NADPH-P450 reductase. The sequence is placed
 CC under control of the yeast ADH gene promoter and terminator. The vectors
 CC are used in a method for evaluating the safety of a cpd. by reacting the
 CC test cpd. with recombinantly produced human cytochrome P450 mol. species
 CC 1A2 (AAT28380), 2C9 (AAT28381), 2E1 (AAT28382), 3A4 (AAT28383) or their
 CC variants (AAT28384-98) together with yeast NADPH-P450 reductase (either
 CC as a fused protein or as a cell extract) and analysing the resultant
 CC metabolite. The cpd. is considered "safe" if it is detoxified or not
 CC rendered carcinogenic or "unsafe" if it is not detoxified or is
 CC metabolised to a carcinogenic cpd
 XX
 SQ Sequence 1494 BP; 248 A; 508 C; 447 G; 291 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,77e-223 Length: 1494
 Score: 2285.50 Matches: 444
 Percent Similarity: 89.54% Conservative: 1
 Best Local Similarity: 89.34% Mismatches: 1
 Query Match: 98.09% Indels: 51
 DB: 2 Gaps: 1
 US-09-820-788A-2 (1-446) x AAT28397 (1-1494)
 QY 1 MetGlyLeuAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
 Db 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCGGTGATAGTGGCCATCTCTCTCTCTG 60
 QY 21 AspLeuMetHisArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40
 Db 61 GACCTTGATCAGCGGCGCCACAGCTGGGTGACGCTACCCACCGAGCCCTTGGCCACTG 120

QY 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
 Db 121 CCCGGGCTGGCAACCTGTGCTGTGGACTTCCAGAACACACATACTGCTTCGACCAG 180
 QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
 Db 181 TTGCGGCGCCGCTTCGGGAGCCTGTTCAGCCTGCAGCTGGCGGCGCGGTGGTGGT 240
 QY 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
 Db 241 CTCATATGGCTGGCGGCGCTGGCGAGGCGCTGGTGACCCACGCGCGAGACACCCGCGAC 300
 QY 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln 117
 Db 301 CGCGCGCCTGTGCCCATCACCATCCTGGGTTCCTGGGCGCGCTTCCCAAGGGGTGTC 360
 QY 117 ----- 117
 Db 361 CTGGCGCGCTATGGGCGCGCTGGCGCGAGAGGCGCTTCTCCGTCTCCACCTTTCGCG 420
 QY 117 ----- 117
 Db 421 AACTTGGGCGCTGGCAAGAGTGGCTGGAGCAGTGGGTGACCGAGAGGCGCGCTGCCTT 480
 QY 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
 Db 481 TGTGGCGCCTTCGCCAACCCACTCCGCGCGCGCTTTCGCCCAACGGTCTCTTGGACAAA 540
 QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspPro 149
 Db 541 GCGGTGAGCAACGTATGCGCTCCCTACCTCGGCGCGCGCTTCGAATACGACGACCCCT 600
 QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
 Db 601 CGCTTCTCGGCTGCTGGACCTAGCTCAGAGGGGACTGAAGAGAGGTGGGCTTTCG 660
 QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
 Db 661 CGCGAGGTGCTGAATGCTGTCCCGTCTCTGCTGCATATCCAGCGCTGGCTGGCAAGGTC 720
 QY 190 LeuArgPheGlnIlyAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
 Db 721 CTACGCTTCCAAAGGCTTCTCAGCCAGCTGGATGAGTGTCTAATCAGCAGCAGGATG 780
 QY 210 ThrTrpAspProAlaGlnProArgPheLeuThrGluAlaPheLeuAlaGluMetGlu 229
 Db 781 ACCTGGGACCCAGCCCGAGCCCTGAGCTGAGGCGCTTCTTGGCAGAGATGGAG 840
 QY 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
 Db 841 AAGGCCAAGGGGAACCTTGAGAGCAGCTTCATGATGAGAACTCGGCATAGTGGTGGCT 900
 QY 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 269
 Db 901 GACCTGTTCTTCGCGGATGGTGACCACTCGACCGCTGGCGCTGGCGCTCTCTGCTC 960
 QY 270 MetIleLeuHisProAspValGlnNargValGlnGlnLeuLeuLeuAspValIleGly 289
 Db 961 ATGATCCTTACATCCCGATGTCAGCGCGCTGTCCAAACAGGAGATCGACACGTGTAGGG 1020
 QY 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
 Db 1021 CAGGTGGCGGACACAGATGGGTGACCGAGGTTCATGTCCTCACCACTCGCGGTGATT 1080
 QY 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
 Db 1081 CATGAGGTGCAGCGCTTTGGGACATCGTCCCTCGGGTGTGACCATATGATCATCCCT 1140
 QY 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
 Db 1141 GACATCGAAGTACAGGGCTTCGCGATCCTTAAGGGAACGACACTCATCAACCACTGTCA 1200

```
QY 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGTGTGTAAGATGAGCCGCTTGGAGAGAGCCCTTCGGCTTCCACCCGAAACACTTC 1260
QY 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGCGCACCTTTGTGAAGCGGAGGCTTCTGCTCTTCTCAGACGCGCG 1320
QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeu 409
Db 1321 CGTGATGCTCGGGAGCGCCCTGCGCGCATGGAGCTTCTCTCTTCTTCACTCCCTG 1380
QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
Db 1381 CTGCAGCACTTCAGCTTCTCGGTGCCACTGGACAGCCCGCGCCAGCACCATGGTGT 1440
QY 430 PheAlaPheLeuValThrProSerProTyTrpGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTTCGTGAGCCATCCCTTATGAGCTTGTGTGTGCGCCGCG 1491
```

RESULT 10

AAT17388
ID AAT17388 standard; cDNA; 1494 BP.

AC AAT17388;

XX 01-AUG-1996 (first entry)

XX Human derived cytochrome P4502D6 cDNA.

DE Human derived cytochrome; P4502D6; commercial cDNA library; yeast;
KW transfection; recombinant production; expression vector; mammal;
KW immunisation; sensitisation; antibody; determination; detection;
KW non-cross reactive; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 1..1494
FT CDS /*tag= a

XX JP08027199-A.

XX 30-JAN-1996.

XX 15-JUL-1994; 94JP-00164186.

XX 15-JUL-1994; 94JP-00164186.

XX (SUMO) SUMITOMO CHEM CO LTD.

XX WPI; 1996-136340/14.

XX P-PSDB; AAR81462.

XX Antibody recognising human derived cytochrome P4502D6 - allows specific
PT detection of cytochrome P450 species in humans.

XX Example 1; Page 11-13; 13pp; Japanese.

XX The present sequence encodes the human derived cytochrome (HDC) P4502D6,
CC which was obtd. from a commercial cDNA library. Yeast were transfected
CC with an expression vector contg. the HDC cDNA, cultured and then
CC disrupted to give a microsomal fraction. The HDC was purified from the
CC fraction, and used to immunise and sensitise a mammal. Blood was drawn
CC from the mammal, and an anti-HDC antibody isolated. The antibody obtd.
CC recognises HDC P4502D6, partic. at a serum dilution rate of 1:10000, and
CC is substantially without cross reaction to other HDC P450 spp

XX Sequence 1494 BP; 248 A; 508 C; 447 G; 291 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.77e-223 Length: 1494
Score: 2285.50 Matches: 444

```
Percent Similarity: 89.54% Conservative: 1  
Best Local Similarity: 89.34% Mismatches: 1  
Query Match: 51  
DB: 2 Indels: 51  
Gaps: 1  
US-09-820-788A-2 (1-446) x AAT17388 (1-1494)  
QY 1 MetGlyLeuGluAlaLeuValProLeuAlaValLleValAlaLlePheLeuLeuVal 20  
Db 1 ATGGGGCTAGAAAGACACTGGTGGCCCTGGCCGTAGTGGCCATCTTCTCTCTCTCTG 60  
QY 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40  
Db 61 GACCTGATGCAACGGCGCAACGCTGGCTGCACGCTACCCACACAGGCCCTCTGCCACTG 120  
QY 41 ProGlyLeuGlyAanLeuLeuHisValAspPheGlnAanThrProTyTrpCysPheAspGln 60  
Db 121 CCCGGGCTGGGCAACCTGCTCATGTGGACTTCAGAACACACCATACTGCTTCGACCAG 180  
QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80  
Db 181 TTGGGGCGCCGCTTGGGGGACGTGTTGAGCTGCGAGCTGGCTGGACGCCGCTGGTGTG 240  
QY 81 LeuAanGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100  
Db 241 CTCAATGGCTGGCGGCGCTGGCGAGCGCTGGTGACCCACGCGGAGGACACACGCCGAC 300  
QY 101 ArgProValProLleThrGlnLeuGlyPheGlyProArgSerGln----- 117  
Db 301 CGCCCGCTGTGCCCATCACCAGATCTGGGTTCGGGGCGCGTTCCCAAGGGGTGTTC 360  
QY 117 ----- 117  
Db 361 CTGGCGGCTATGGGCGCGCTGGCGAGCAGAGCGCTTCTCGTCTCCACCTTGGCGC 420  
QY 117 ----- 117  
Db 421 AACTTGGGCTGGGCAAGAGTCCGTGGAGCAGTGGGTGACCGAGGAGCGCGCTT 480  
QY 118 -----GlyArgProPheArgProAanGlyLeuLeuAspLys 129  
Db 481 TGTGCGGCTTTCGCCAACCACTCCGGAGCGCCCTTTGGCCCAACGGTCTCTTGGACAA 540  
QY 130 AlaValSerAanValLleAlaSerLeuThrCysGlyArgArgPheGluTyTrpAspPro 149  
Db 541 GCCGTGAGCAACGTGATGCGCTCCCTCACCCTGGGGCGCGCTTCGAATACGACACCT 600  
QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169  
Db 601 CGCTTCTCAGGCTGCTGGACCTAGCTCAGAGGGGACTGAGAGGAGGAGTCCGGCTTCTG 660  
QY 170 ArgGluValLeuAanAlaValProValLeuLeuHisLleProAlaLeuAlaGlyLysVal 189  
Db 661 CGCAGGTGCTGAATGCTGCTCCCGTCTCTGATATATCCAGCGCTGGTGGCAAGGTC 720  
QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209  
Db 721 CTACGCTTCCAAAAGGCTTTCCTACCCAGCTGGATGAGCTGCTAACTGACACAGGATG 780  
QY 210 ThrTrpAspProAlaGlnProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229  
Db 781 ACCTGGGACCCAGCCAGCCCGCCCGACCTGACTGAGGCTTCTTGGCAGAGATGGAG 840  
QY 230 LysAlaLysGlyAanProGluSerPheAanAspGluAanLeuArgIleValValAla 249  
Db 841 AAGCCAAAGGGGAACCTCGAGCAGCTTCAATGATGAGAACCTCGCGCATAGTGGTGGCT 900  
QY 250 AspLeuPheSerAlaGlyMetValThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 269  
Db 901 GACCTGTCTCTGCGGGATGGTACCACTTCGACCACTCGTGGCGCTGGGGCTCTCTGCTC 960  
QY 270 MetIleLeuHisProAspValGlnArgValGlnGlnLeuLeuAspValIleGly 289
```

```
Db 961 ATGATCCTACATCCGATGTGACGCCGCTGTCCAAAGAGAGATCGACGACGTGATAGGG 1020
Qy 290 GlnValArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
Db 1021 CAGTGGCGGACACAGAGATGGGTGACAGGCTCACATGCCCTACACCACTGCCGTGATT 1080
Qy 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATGAGGTGCAGCGCTTTGGGGACATCGTCCCGCTGGGTGTGACCATATGACATCCCGT 1140
Qy 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAGTACAGGGCTTCGCATCCCTAAGGGAAACGACATCATCAACCACTGTCA 1200
Qy 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAAGCATCAGCGCGTCTGGAGAGCGCTTCGGCTTCCACCCGCAACTTC 1260
Qy 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCCGAGGCGCACTTTGTGAAGCGGAGCGCTTCCTGCCTTTCTCAGCAGCGCC 1320
Qy 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 409
Db 1321 CGTGCATGCTCGGGAGCGCCCTGGCGCGCATGGAGCTCTTCTCTTCTTCACTCCCTCG 1380
Qy 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
Db 1381 CTGCAGCACTTACGCTTCTCGGTGCCCACTGGACAGCCCGCGCCAGCCACCATGTGTGTC 1440
Qy 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1441 TTGTCTTCTCGTGAGCCCATCCCGCTATGAGCTTTGTGTGTGCGCCGCGC 1491

RESULT 11
AB072216
ID ABQ72216 standard; cDNA; 1494 BP.
XX
AC ABQ72216;
XX
DT
XX
XX 02-SEP-2002 (first entry)
DE Human CYP2D6 gene coding sequence, SEQ ID NO:2.
XX
KW Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;
KW chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;
KW antiarrhythmic; arrhythmia; adrenoreceptor antagonist; hypertension;
KW tricyclic antidepressant; procainamide; drug induced lupus syndrome;
KW environmentally linked disease; Parkinson's disease; haplotyping;
KW genotyping; haplotype; genetic variant; single nucleotide polymorphism;
KW SNP; drug screening; drug discovery; gene; ss.
XX
OS Homo sapiens.
XX
FH
FH Key Location/Qualifiers
FH CDS 1..1494
FT /*tag= a
FT /product= "CYP2D6"
FT /replace(19, A)
FT /label= PS7
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V7M"
FT /tag= c
FT /label= PS8
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V11W"
FT /tag= d
FT /label= PS9
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution P34S"
FT

replace(263, A)
/*tag= e
/*label= PS12
/*note= "Novel single nucleotide polymorphism (SNP);
causes the amino acid substitution R88H"
replace(271, A)
/*tag= f
/*label= PS13
/*note= "Known single nucleotide polymorphism (SNP);
causes the amino acid substitution L91M"
replace(281, G)
/*tag= g
/*label= PS14
/*note= "Novel single nucleotide polymorphism (SNP);
causes the amino acid substitution H94R"
replace(294, G)
/*tag= h
/*label= PS15
/*note= "Novel single nucleotide polymorphism (SNP)"
replace(311, C)
/*tag= i
/*label= PS16
/*note= "Novel single nucleotide polymorphism (SNP);
causes the amino acid substitution V104A"
replace(319, T)
/*tag= j
/*label= PS17
/*note= "Novel single nucleotide polymorphism (SNP);
together with PS18 causes the amino acid substitution
T107P"
replace(320, T)
/*tag= k
/*label= PS18
/*note= "Novel single nucleotide polymorphism (SNP);
together with PS17 causes the amino acid substitution
T107P"
replace(325, G)
/*tag= l
/*label= PS19
/*note= "Novel single nucleotide polymorphism (SNP);
causes the amino acid substitution I109V"
replace(333, C)
/*tag= m
/*label= PS20
/*note= "Novel single nucleotide polymorphism (SNP)"
replace(336, T)
/*tag= n
/*label= PS21
/*note= "Known single nucleotide polymorphism (SNP)"
replace(358, A)
/*tag= o
/*label= PS27
/*note= "Novel single nucleotide polymorphism (SNP);
causes the amino acid substitution F120I"
replace(382, C)
/*tag= p
/*label= PS28
/*note= "Novel single nucleotide polymorphism (SNP);
causes the amino acid substitution W128R"
replace(406, A)
/*tag= q
/*label= PS29
/*note= "Novel single nucleotide polymorphism (SNP);
together with PS30 causes the amino acid substitution
V136I"
replace(408, C)
/*tag= r
/*label= PS30
/*note= "Known single nucleotide polymorphism (SNP);
together with PS29 causes the amino acid substitution
V136I"
replace(451, G)
/*tag= s
```

FT /label= PS31
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution Q151E"
FT replace(463, A)
FT /*tag= t
FT /label= PS32
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution E155K"
FT replace(696, C)
FT /*tag= u
FT /label= PS35
FT /note= "Known single nucleotide polymorphism (SNP)"
FT replace(1012, A)
FT /*tag= v
FT /label= PS36
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V338M"
FT replace(1030, T)
FT /*tag= w
FT /label= PS37
FT /note= "Novel single nucleotide polymorphism (SNP);
FT generates a premature stop codon (R344STOP)"
FT replace(1083, C)
FT /*tag= x
FT /label= PS38
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace(1457, C)
FT /*tag= y
FT /label= PS41
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution S466T"
FT replace(1489, T)
FT /*tag= z
FT /label= PS42
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution R497C"
XX WO200239589-A2.
XX 16-MAY-2002.
XX 09-NOV-2001; 2001WO-US047396.
XX 09-NOV-2000; 2000US-0247943P.
XX (GENA-) GENAISSANCE PHARM INC.
XX Anastasio AE, Chew A, Choi JY, Denton RR, Nandabalan K;
XX Petersen N, Rounds E;
XX WPI; 2002-519292/55.
XX P-PSDB; AB09563.
XX Novel genetic variants of Cytochrome P450, Subfamily IID, Polypeptide 6
XX isoenzymes, useful for improving efficiency and reliability in drug
XX development for treating hypertension, arrhythmias and Parkinson's
XX disease.
XX Claim 25; Fig 2; 158pp; English.
XX The invention relates to a method for haplotyping the cytochrome P450,
XX subfamily IID, polypeptide 6 (CYP2D6) gene (ABQ72215, ABQ72364) of an
XX individual, and also describes 29 novel polymorphic sites within the
XX human CYP2D6 gene. The CYP2D6 gene is located on chromosome 22q13.1 and
XX contains 9 exons which encode a 497 amino acid protein (AB09563). CYP2D6
XX is a mono-oxygenase involved in the detoxification of many drugs and
XX environmental chemicals. It plays a role in the metabolism of drugs such
XX as antiarrhythmics, adrenoceptor antagonists and tricyclic
XX antidepressants, and is also involved in the formation of a metabolite
XX linked to the drug-induced lupus syndrome observed with procainamide.
XX Variations in CYP2D6 activity or expression may also influence an
XX individual's susceptibility to environmentally-linked diseases, and it
XX has been demonstrated that CYP2D6 activity may be involved in the

CC pathogenesis of Parkinson's disease, with individuals with a less active
CC form of the enzyme tending to have an earlier onset of this condition.
CC CYP2D6 nucleic acid sequences are useful in studying the expression and
CC function of CYP2D6, and in expressing CYP2D6 protein for use in screening
CC drugs for the treatment of CYP2D6-associated diseases (e.g.,
CC hypertension, atrial and ventricular arrhythmias, Parkinson's disease,
CC and drug-induced lupus syndrome) or which are metabolised by CYP2D6.
CC CYP2D6 nucleic acids and proteins are also useful in studying the effect
CC of polymorphisms on the biological activity of CYP2D6. Polymorphisms in
CC the target region may be determined by the use of allele-specific
CC oligonucleotides (ASOs; ABQ72217-ABQ72303) as probes and primers, and by
CC primer extension using oligonucleotide primers comprising sequences
CC ABQ72304-ABQ72361. The method of the invention is useful for haplotyping
CC the CYP2D6 gene in populations and in individuals, enabling decisions to
CC be made as to whether CYP2D6 is a likely therapeutic target for a disease
CC of interest, and to control for genetically-based bias in the design of
CC drugs that target or are metabolised by CYP2D6. In addition, transgenic

Alignment Scores:
Pred. No.: 1,77e-223 Length: 1494
Score: 2285.50 Matches: 444
Percent Similarity: 89.54% Conservatives: 1
Best Local Similarity: 89.34% Mismatches: 1
Query Match: 98.09% Indels: 51
DB: 6 Gaps: 1

US-09-820-788A-2 (1-446) x ABQ72216 (1-1494)

QY 1 MetGlyLeuGluAlaLeuValProLeuAlaValLeuValAlaLeuPheLeuLeuVal 20
DB 1 ATGGGGCTAGAGACACTGGTGGCCCTGGCGCTGATAGTGGCCATCTTCTCTCTCTG 60

QY 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeu 40
DB 61 GACCTGATGCACCGCGCCAAACGCTGGCTGCACGCTACCCACGAGGCCCTGCACTG 120

QY 41 ProGlyLeuGlyAsnLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
DB 121 CCGGGCTGGGCAACCTGCTGATGTGACTTCAGAACACACCATCTGTTGACCCAG 180

QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValVal 80
DB 181 TTGGGGCGCGCTTCGGGACGCTGTTCAAGCTGCGAGCTGGCGTGGACCGCGTGGT 240

QY 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
DB 241 CTCAATGGGCTGGCGCGCGCTGGCGAGCGCTGGTGACCCACGCGAGGACACCGCCGAC 300

QY 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln 117
DB 301 CGCCGCGCTGTGCCCATCACCAGATCTCTGGGTTCGGGGCGCGGTTCACAGGGGTG 360

QY 117 ----- 117
DB 361 CTGGCGCGCTATGGCGCGCGCTGGCGGAGCAGAGCGCTTCTCGTGTCCACCTTGGCG 420

QY 117 ----- 117
DB 421 AACTTGGGCTGGGCAAGAGTCGCTGGAGCAGTGGGTGACCGAGGAGCGCGCTT 480

QY 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
DB 481 TGTGGCGCTTGGCGAACCACTCGGAGCGCCCTTTCGCCCAACGGTCTCTTGGACAAA 540

QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGlyTyrAspAspPro 149
DB 541 GCGGTGAGCAAGTGTGCTCCCTCCTACCTGCGGGCGCGCTTCGAGTACGACGACCT 600

QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluLysArgGlyPheLeu 169
DB 601 CGCTTCTCAGCTGCTGAGCTAGCTCAGGAGGAGGAGTGAAGAGGAGTTCGGCTTCTG 660

QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189

```
Db 661 CGCGAGGTCGTAATCTGCTCCCGCTCTCCATATCCAGCGCTGGCTGGCAAGGTC 720
Qy 190 LeuAtpPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuThrGluHisArgMet 209
Db 721 CTACGGCTTCAAAGAGCTTCTGACCCAGCTGGATGAGCTGCTAACTGAGCACAGGATG 780
Qy 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 781 ACTGGGACCCAGCCCGAGCCCGGAGACCTGACTGAGGCTTCTTGGCAGAGATGGAG 840
Qy 230 LysAlaLysGlyAsnProGlnSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
Db 841 AAGGCCAAGGGAACCTCAGAGACGCTTCAATGATGAGAACCTGCGCATAGTGGTGGCT 900
Qy 250 AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTrpGlyLeuLeuLeu 269
Db 901 GACCTGTCTCTGCGGGATGGTGACCACTGACCACTGCGCTGGCGCTCTCTGCTC 960
Qy 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnIleAspAspValIleGly 289
Db 961 ATGATCTACATCCGATGTCAGCGCGCTGTCCAAACAGAGATCGACGAGTATAGG 1020
Qy 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValle 309
Db 1021 CAGGTGCGCGGACCAAGATGGGTGACCAAGGCTCACATGCCCTACCACTCCGCTGATT 1080
Qy 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATGAGTGCAGCGCTTTCGGGACATCGTCCCTCGGTGGTGACCCATATGATCCCT 1140
Qy 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAAGTACAGGCTTTCGCGATCCCTAAGGGAACGACACTCATCAACCTGTCA 1200
Qy 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAAGGATGAGCGCTCTGGGAGAGCCCTTCGGCTTCCACCCCGAACCTTC 1260
Qy 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGCGCACCTTTGTGAAGCGGAGGCGCTTCTGCTTTCTCAGACGCGCG 1320
Qy 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeu 409
Db 1321 CGTGATGCTCTGGGAGCGCTGGCGCGCATGGAGCTTCTCTTCTTCACTCCCTG 1380
Qy 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
Db 1381 CTGCAGCACTTCAGCTTCTCGGTGCGCCACTGGACAGCGCGCGCCAGCCACCATGGTGC 1440
Qy 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1441 TTGTGTTTCTGTGTAGCCCATCCCCCTATGAGCTTTGTGTGTGCGCGCG 1491
```

RESULT 12

ADB25778

ID ADB25778 standard; DNA; 1567 BP.

XX ADB25778;

AC ADB25778;

XX 20-NOV-2003 (first entry)

DT Human CYP2D6-related gene #3.

XX human; mutant CYP2D6 gene; drug analysis; drug testing; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT 1..1494

FT /tag= a

FT /product= "Human CYP2D6-related protein #3"

FT

```
XX WO2003050282-A1.
PN 19-JUN-2003.
XX 05-DEC-2002; 2002WO-JP012748.
XX 06-DEC-2001; 2001JP-00372548.
XX (TSUR ) TSUMURA & CO.
XX Taniyama M, Ogawa K, Tsuchiya N, Hibino T;
XX WPI; 2003-505401/47.
XX P-PSDB; ADB25833.
XX Genetic polymorphisms of CYP2D6 gene in human population for analysis of
XX drug effect on individual patients and testing of new drugs.
XX Claim 8; Page 46-50; 75pp; Japanese.
XX The invention comprises mutant forms of the human CYP2D6 gene, containing
XX one or more of the following mutations G125A, C1858T, T2874C and C2875T.
XX The mutant human CYP2D6 genes of the invention are useful for analyzing
XX the effect of drugs on individual patients and testing of new drugs. The
XX present DNA sequence represents a human gene of the invention.
XX SQ Sequence 1567 BP; 266 A; 528 C; 465 G; 308 T; 0 U; 0 Other;
```

Alignment Scores:

```
Pred. No.: 1.9e-223 Length: 1567
Score: 2285.50 Matches: 444
Percent Similarity: 89.54% Conservative: 1
Best Local Similarity: 89.34% Mismatches: 1
Query Match: 98.09% Indels: 51
DB: 7 Gaps: 1
```

US-09-820-788A-2 (1-446) x ADB25778 (1-1567)

```
Qy 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
Db 1 ATGGGGCTAGAGACACTGGTCCCTGGCGGTAGTAGGCGCATCTTCTGCTCTCTGGTG 60
Qy 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40
Db 61 GACCTGATGCACCGCGCGCAACGCTGGCTGGCTGACCTACCCACAGGCCCTTGCACCTG 120
Qy 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 121 CCGCGGCTGGGCAACCTGCTGCATGTGGACTTCCAGAAACACACCATACTGCTTCCAGCAG 180
Qy 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
Db 181 TTGGCGCGCGCTTCCGGGAGCGTGTTCAGCTGACCTGGAGCTGGCGCGCGGTGGTGGTG 240
Qy 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTCATGGCTGGCGCGCGCTGGCGCGCGCTGGTGACCCACGCGGAGGACACCGCGGAC 300
Qy 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln----- 117
Db 301 CGCGCGCGCTGTGCCCATCACCAGATCCTGGGTTTCGGGCGCGCTTCCAGAGGGGTGTTTC 360
Qy 117 ----- 117
Db 361 CTGGCGCGCTATGGCGCGCGCTGGCGCGAGAGAGCGCTTCTCCGTCTCCACCTTGGCGC 420
Qy 117 ----- 117
Db 421 AACTTGGCGCTGGGCAAGAGTCTGGAGCAGTGGTGACCGAGAGGCGCGCTTGCCTT 480
Qy 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
Db -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
```

```
Db 481 TGTGGCGCTTCGCCAACACCTCCGAGCGCCCTTTGGCCCAACGGTCTCTTGGACAAA 540
Qy 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyraAspPro 149
Db 541 GCCGTGAGCAAGTATGCCCTCCCTACCTCGCGCGCGCGCTTCGAGTACACACCT 600
Qy 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLeuGluSerGlyPheLeu 169
Db 601 CGCTTCCTCAGGCTGCTGACCTAGCTCAGGAGGACCTGAAGAGGAGTCCGGCTTCG 660
Qy 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
Db 661 CGCGAGGTGCTCAATGCTGTCCTCCCTCTCTGCATATCCAGCGCTGCTGCAAGTCT 720
Qy 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
Db 721 CTACGCTTCCAAAGGCTTCTCTGACCCAGCTGAGTGTCTAACTGAGCAGCAGGATG 780
Qy 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 781 ACTTGGACCCAGCCAGCCCGGACCTGACTGAGGCTTCTTGGCAGAGATGGAG 840
Qy 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
Db 841 AAGGCCAAGGGAAACCTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGTGCT 900
Qy 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeu 269
Db 901 GACCTGTCTCTGCGGGATGTGACCACTGACCTGACCTGCGCTGCGGCTCTCTGCTC 960
Qy 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluLeuAspValIleGly 289
Db 961 ATGATCTTACATCCGATGTGCGAGCGCTGTCACACAGGAGATCGACAGCTGATAGG 1020
Qy 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyThrThrAlaValIle 309
Db 1021 CAGGTGCGCGACACAGAGATGGTGACCAAGCTCATGCTCCCTACCACTGCGGTGAT 1080
Qy 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATGAGGTGACGCTTTGGGACATCGTCCCTCGGTGTGACCATATGACATCCCGT 1140
Qy 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAAGTACAGGCTTCGCACTCCCTAAGGAACGACACTCATCACCACTGTCA 1200
Qy 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAGGATGAGCGCTCTGGGAGAGCCCTTCGCTTCCACCCGAACTTC 1260
Qy 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGCGCACTTTGTGAAGCGGAGGCTTCCTGCTCTCTGACGAGCGCG 1320
Qy 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 409
Db 1321 CGTGATGCTCGGGGAGCCCTGCGCGCATGGAGCTTCTCTCTTCTTACCTCCCTG 1380
Qy 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 429
Db 1381 CTGCAGCACTTCAGCTTCTGGTGGCCCACTGGAGACCCCGGCGCCAGCCACCATGTGTC 1440
Qy 430 PheAlaPheLeuValThrProSerProTyThrGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTTCTGCTGAGGCCCATCCCTCATGAGCTTTGTGCTGTGCTGCTGCTGCTG 1491
RESULT 13
ACF06056
ID ACF06056 standard; cDNA; 1494 BP.
XX
AC ACF06056;
XX
DT 22-SEP-2003 (first entry)
```

```
XX Human cytochrome P450 2D6 encoding cDNA.
DE
XX Human; protein array; protein moiety; phenotype; drug discovery;
KW naturally occurring variant; pharmacogenomic; diagnostic;
KW parallel analysis; tumour suppressor; p53; cytochrome P450; gene; ss.
XX Homo sapiens.
XX WO2003048768-A2.
PN 12-JUN-2003.
XX
XX 05-DEC-2002; 2002WO-GB005499.
XX
XX 05-DEC-2001; 2001US-0335806P.
PR 16-SEP-2002; 2002US-0410815P.
XX
XX (SENS-) SENSE PROTEOMIC LTD.
XX
XX Boutell JM, Godber BLJ, Hart DJ, Blackburn JD;
PI WPI; 2003-569063/53.
XX P-PSDB; ABR82026.
XX
XX New protein array, useful for determining the phenotype of a naturally
PT occurring variant of a DNA sequence of interest, comprises a surface upon
PT which at least two protein moieties are deposited.
XX
XX Example 5; Fig 13A; 84pp; English.
XX
XX The present invention describes a protein array comprising a surface upon
CC which at least two protein moieties are deposited at spatially defined
CC locations, where the protein moieties are naturally occurring variants of
CC a DNA sequence of interest. Also described: (1) making a protein array;
CC (2) screening a set of protein moieties for molecules that interact with
CC one or more proteins; and (3) simultaneously determining the relative
CC properties of members of a set of protein moieties. The protein array can
CC be used for determining the phenotype of a naturally occurring variant of
CC a DNA sequence of interest. The protein array is useful for drug
CC discovery, pharmacogenomics and diagnostics. The protein array allows the
CC parallel analysis of closely related proteins with a sensitivity that is
CC at least comparable to existing methods, if not better, with small
CC volumes of potentially expensive ligands, and in a quantitative,
CC comparative functional analysis manner not previously possible. ACF06000
CC to ACF06056 and ABR81975 to ABR82026 represent sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 1494 BP; 248 A; 507 C; 448 G; 291 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.59e-223 Length: 1494
Score: 2282.50 Matches: 443
Percent Similarity: 89.54% Conservative: 2
Best Local Similarity: 89.13% Mismatches: 1
Query Match: 97.96% Indels: 51
DB: 8 Gaps: 1
US-09-820-788A-2 (1-446) x ACF06056 (1-1494)
Qy 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
Db 1 ATGGGGCTAGAAAGCACTGGTCCCTCGCGGTGATGGCCATCTTCTCTCTCTCTG 60
Qy 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40
Db 61 GACCTGATGACCGCGCAACGCTGGGCTGACGCTACCCACGAGGCCCTTGCACCTG 120
Qy 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyThrCysPheAspGln 60
Db 121 CCCGGGCTGGGCAACCTGCTGCTGACATTCAGCAACACACCATATCTGCTTGCACGAG 180
Qy 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
```


181 TTGCGCGCGCTTGGGAGCGTGTTCAGCTCGAGCTGGCTGGACGCGGTGTGCTG 240
81 LeuAnGlyLeuAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
241 CTCATGGGCTGGCGCGCTGGCGAGGCGCTGTGACCCACGCGAGACACCCCGAC 300
101 ArgProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln 117
301 CGCCGCGCTGTGCCCATCACCCAGATCCTGGGTTCCTGGGCGCGCTTCCCAAGGGGTGTTTC 360
117 ----- 117
361 CTGGCGCGCTATGGGCGCGCTGGCGGAGAGAGCGCTTCTCGGTGTCCACCTTGGCG 420
117 ----- 117
421 AACTTGGGCTGGCGAGAGTGGCTGGAGAGTGGGTGACCGAGAGCGCCCGCTTGCCTT 480
118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
481 TGTGGCGCTTCGCCAACCACTCCGGAGCGCCCTTTCGCCCAACGGTCTCTTGGACAA 540
130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluThrAspAspPro 149
541 GCGGTGAGCAACGTCATCGCTCCCTCACCTCGGGCGCGCTTCGAGTACGACGACCT 600
150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
601 CGCTTCCTCAGGCTGTGACCTAGCTCAGGAGGAGCTGAAGAGGAGTGGGGCTTTCG 660
170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
661 CGCGAGTGTCTGAATGCTCTCCCGCTCTCTCTGCATATCCAGCGCTGGCTGCGAAGTTC 720
190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
721 CTACGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGCTGAATGAGCACTGAGCAGGATG 780
210 ThrTroAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
781 ACCTGGACCCAGCCAGCCCGGAGACCTGACTGAGGCGCTTCTGGCAGAGATGGAG 840
230 LysAlaLysGlyAsnProGluSerPheAsnAspGluAsnLeuArgIleValValAla 249
841 AAGGCCAAGGGACCTCAGAGCAGCTTCATGATGAGAACTGCGCATAGTGTGGCT 900
250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 269
901 GACCTGTTCTTGGCGGATGGTGACCACTCGACACGCTGGCGCTGGCGCTTCTGTCTC 960
270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspAspValIleGly 289
961 ATGATCTCATCTCGGATGTGAGCGCGCTGTCCAAACAGGAGATCGACGACGTGATGGG 1020
290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
1021 CAGGTGGCGGACGAGATGGGTGACAGGCTCATGCTCCCTACCACTCCCGGTGATT 1080
310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
1081 CATGAGTCCAGCGCTTTCGGGACATCGTCCCGCTGGGTATGACCAATATGACATCCCGT 1140
330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
1141 GACATCGAAGTACAGGCTTCGCGATCCCTTAAGGGAACACACTCATCAACCACTGTCA 1200
350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
1201 TCGGTCTGAAGATAGCGCGTCTGGGAGAGCCCTTCCGTTCCACCCCGAAGCACTTC 1260
370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389

1261 CTGATGCCCGAGGGCCACTTTGTGAAGCCGAGGCGCTTCTCGCTTTCTCAGCAGGCGCG 1320
390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrSerLeu 409
1321 CGTGCATGCTCGGGAGCCCTGGCCGCGCATGGAGCTTCTCTTCTTCCACTCCCTG 1380
410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
1381 CTGCAGCACTTTCAGCTTCTCGGTGCCACTGACAGCCCGCGCCAGCACCACCATGGTGTTC 1440
430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
1441 TTTGCTTTCTGGTGAGCCCATCCCTTATGAGCTTTGTGTGTGCCCCGC 1491
RESULT 14
ACA61303
ID ACA61303 standard; cDNA; 1567 BP.
XX
AC ACA61303;
XX
DT 16-JUL-2003 (first entry)
XX
DE Human cytochrome p450 gene CYP2D6, wild-type cDNA.
XX
KW Human; ss; gene; cytochrome P450; CYP2D6; chromosome 22; SNP;
KW single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
KW psychiatric disorder; drug sensitivity.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1494
FT FT /*tag= a
FT FT /product= "CYP2D6 protein"
XX
PN EP1281755-A2.
XX
PD 05-FEB-2003.
XX
PF 16-JUL-2002; 2002EP-00254972.
XX
PR 31-JUL-2001; 2001US-0309111P.
XX (PFIZ) PFIZER PROD INC.
XX
PI Milos PM, Webb SM;
XX
DR WPI; 2003-373769/36.
DR P-PSDB; ABU09593.
XX
PT New cytochrome P450 2D6 gene variants and polypeptides, useful for
PT determining if a subject has or is at risk of developing a drug
PT sensitivity condition or disorder that is associated with an aberrant
PT CYP2D6 activity.
XX
PS Claim 2; Fig 4; 88pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising a cytochrome
CC P450 2D6 gene variant, e.g. G5799C or C5816A (referring to the genomic
CC sequence or the same variant nucleotide in the corresponding cDNA
CC sequences). Also included are probes, primers (allele specific
CC oligonucleotides) and arrays used to detect and/or amplify the CYP2D6
CC gene polymorphic regions, the variant polypeptides, antibodies which are
CC capable of distinguishing between the variant and wild-type polypeptides,
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
CC determining whether an individual is susceptible to being a poor
CC metaboliser of drugs. The DNA probe is useful for hybridising to a
CC variant form of the CYP2D6 gene. The primer is useful for amplifying the
CC C5816A allelic variant. The allele specific nucleotide is useful for the
CC detection of the C5816A allelic variant. The methods are useful for
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and

CC determining if an individual is susceptible to being a poor metaboliser
 CC of drugs. The nucleic acids are useful as probes or primers for
 CC determining whether a subject has a genetic deficiency for metabolising
 CC drugs that are substrates of P450 CYP2D6. The methods are useful for
 CC determining if a subject has or is at risk of developing a drug
 CC sensitivity condition or disorder that is associated with an aberrant
 CC CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
 CC aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
 CC appropriate drugs or determining the course of treatment to administer to
 CC a subject to treat cardiovascular or psychiatric disorders, or for
 CC treating a subject with a drug sensitivity or disorder associated with a
 CC specific allelic variant of a polymorphic region of the CYP2D6 gene. The
 CC antibodies are useful for monitoring CYP2D6 protein levels in an
 CC individual for determining whether a subject has a disease or conditions
 CC associated with an aberrant CYP2D6 protein level. The gene is located on
 CC human chromosome 22. The present sequence is the wild-type CYP2D6 cDNA
 XX

SQ Sequence 1567 BP; 267 A; 527 C; 465 G; 308 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3,848-223	Length:	1567
Score:	2282.50	Matches:	443
Percent Similarity:	89.54%	Conservative:	2
Best Local Similarity:	89.13%	Mismatches:	1
Query Match:	97.96%	Indels:	51
		Gaps:	1

US-09-820-788a-2 (1-446) x ACA61303 (1-1567)

Qy	1	MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal	20
Db	1	ATGGGGCTAGAGCACTGGTCCCTGGCCGTGATAGTGGCCATCTTCTGCTCTGCTG	60
Qy	21	AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu	40
Db	61	GACCTGATGACCCGGCGCCCAACGGCTGGCGCTGCACGCTACCCACAGGCGCCCTG	120
Qy	41	ProGlyLeuGlyAsnLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln	60
Db	121	CCCGGGCTGGGCACTGCTGATGTGACTTCAGACACACACACACACACACACAC	180
Qy	61	LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaIleThrProValVal	80
Db	181	TTGGCGCGCGCTTCGGGGACGTGTTACGCTGCAGCTGGCGTGGCGCGCGTGGT	240
Qy	81	LeuAsnGlyLeuAlaAlaValArgGlnAlaLeuValThrHisGlyAspThrAlaAsp	100
Db	241	CTCAATGGGCTGGCGCGCGTGGCGAGCGCTGGTGACCCACGCGGAGGACACCG	300
Qy	101	ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln	117
Db	301	CGCCCGCGCTGTGCCCATCACCAAGATCTGGGTTTCGGGGCGCGCTTCCAAAGGG	360
Qy	117	-----	117
Db	361	CTGGCGCGCTATGGSCCGCGTGGCGGAGCAGAGCGCTTCTCGTGTCCACCTTGG	420
Qy	117	-----	117
Db	421	AACTTGGCGCTGGGCAAGAGTCGCTGGAGCAGTGGGTGACCGGAGCGCGCTGCTT	480
Qy	118	-----GlyArgProPheArgProAsnGlyLeuLeuAspLys	129
Db	481	TGTGGCGCTTGGCCAAACCATCCGCGAGCGCCCTTTCGCCCAACGGTCTTTGG	540
Qy	130	AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGlyTyrAspAsp	149
Db	541	GGCGTGACACAGTATGCTCTCCCTACCTGGCGGCGCGCTTCGAGTACGACGACCT	600
Qy	150	ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu	169
Db	601	CGCTTCTCAGCGCTGCTGAGCTAGCTCAGGAGGAGTGAAGGAGGAGTGGGCTTCTG	660

Qy	170	ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal	189
Db	661	CGCGAGGTGCTGAATGCTGTCCCGTCTCTCTGCATATCCCGAGCGTGGCTGGCAAGGTC	720
Qy	190	LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet	209
Db	721	CTAGCGTTCCAAAAGGCTTCTGCTACCCAGCTGGATGAGCTGCTAACTGAGCACAGGATG	780
Qy	210	ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu	229
Db	781	ACCTGGGACCCAGAGCCCGAGCCCGAGACCTGAGTGGGCTTCTCTGGCAGAGATGGAG	840
Qy	230	LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla	249
Db	841	AAGGCCAAGGGGAACCTTGAGAGCAGCTTCAATGATGAGAACCTTGGCGATAGTGGT	900
Qy	250	AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaIleThrGlyLeuLeu	269
Db	901	GACCTGTCTCTGCGGGATGGTGCACCACTCGACCACTGGCTGGCGCTCTCTGCTC	960
Qy	270	MetIleLeuHisProAspValGlnArgArgValGlnGlnGlnIleAspAspValIleGly	289
Db	961	ATGATCTCTACATCCGGATGTGCGCGCGCTGTCCAACAGGAGATCGACGCTGATAGGG	1020
Qy	290	GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle	309
Db	1021	CAGGTGGCGCACAGAGATGGTGCACGAGCTCACATGCCCTACACACTGCGCTGATT	1080
Qy	310	HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg	329
Db	1081	CATGAGGTGCAGCGCTTTTGGGGACATCGTCCCTCGGTGGTATGACCAATATGACATCC	1140
Qy	330	AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer	349
Db	1141	GACATCGAAGTACAGGGCTTCCGCATCCCTAAGGGAACGACACTCATCACCAACCTGTC	1200
Qy	350	SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe	369
Db	1201	TCGGTGTGAAGATGAGCGCTCTGGGAGAGCGCTTCCGCTTCCACCCCGAACACTTC	1260
Qy	370	LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg	389
Db	1261	CTGATGCCAGGCGCACTTTGTGAAGCGGAGCGCTTCTGCTGCTTCTCAGCAGGCGCG	1320
Qy	390	ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu	409
Db	1321	CGTGCATGCTTCGGGGAGCGCTGCGCGCGCATGGAGCTTCTCTCTTCTTCCCTCCCTG	1380
Qy	410	LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal	429
Db	1381	CTGAGCATTTTCAGCTTCTCGGTGCGCCACTGGAGAGCGCCCGCCAGCCCGCCAGCC	1440
Qy	430	PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg	446
Db	1441	TTTCTTCTTCTGTTGAGCCCATCCCTCTATGAGCTTTGTGCTGTGCCCCGCG	1491
XX	RESULT 15		
XX	AAQ87729		
XX	ID	AAQ87729	standard; cDNA; 1494 BP.
XX	AC	AAQ87729;	
XX	XX		
XX	DT	25-MAR-2003	(revised)
XX	DT	15-NOV-1995	(first entry)
XX	XX		Human auxillary cytochrome P450 species 2D6 coding region.
XX	DE		
XX	KW		Human cytochrome P450; amplification; PCR; primer; expression vector;
XX	KW		yeast NADPH-P450 reductase; safety; fusion protein; metabolite;
XX	XX		carcinogen; mutagen; liver metabolism; ds.
XX	OS		Homo sapiens.
XX	XX		

PN EP644267-A2.
 XX 22-MAR-1995.
 XX 20-JUL-1994; 94BP-00111298.
 XX 20-JUL-1993; 93JP-00201120.
 PR 21-JUL-1993; 93JP-00180246.
 PR 30-JUL-1993; 93JP-00208279.
 XX (SUMO) SUMITOMO CHEM CO LTD.
 XX Hayaashi K, Sakaki T, Yabusaki Y, Komai K, Kaneko H, Nakatsuka I;
 PI WPI; 1995-116991/16.
 DR P-PSDB; AAR72375.
 DR Evaluation of safety of a chemical cpd. - using recombinant yeast
 XX expressing human cytochrome p450 and a yeast NADPH-P450 reductase.
 PT Example; Page 82-84; 124pp; English.
 XX The nucleotide sequence of the cDNA coding region for the human auxiliary
 CC cytochrome P450 species 2C6. The gene encodes a protein of 497 amino
 CC acids. The cDNA was amplified by PCR using the primers AAQ87763-6. The
 CC product was cloned into the yeast expression vectors pAAH5N or pAHR5 to
 CC produce the vectors p2D6 for the expression of the cytochrome P450 alone
 CC or p2D6R for co-expression with the yeast NADPH-P450 reductase. The
 CC vectors are used in a method for evaluating the safety of a chemical
 CC compound by reacting the chemical compound with recombinantly produced
 CC human cytochrome P450 molecular species 1A2 (AAQ87714), 2C9 (AAQ87715),
 CC 2E1 (AAQ87716), or 3A4 (AAQ87717) or their auxiliary species and variants
 CC (AAQ87718-32), and yeast NADPH-P450 reductase, either as a fused protein
 CC or in cell extracts, and analysing the resulting metabolite to assess the
 CC safety of the chemical compound. The method is useful for determining
 CC whether the chemical compound, or its metabolite, will be converted into
 CC a carcinogenic or mutagenic form through metabolism in the liver.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 1494 BP; 248 A; 508 C; 446 G; 292 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 4,548-223 Length: 1494
 Score: 2281.50 Matches: 444
 Percent Similarity: 89.34% Conservative: 0
 Best Local Similarity: 89.34% Mismatches: 2
 Query Match: 97.92% Indels: 51
 DB: 2 Gaps: 1

US-09-820-788A-2 (1-446) x AAQ87729 (1-1494)
 QY 1 MetGlyLeuAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
 DB 1 ATGGGGCTAGAGCAGCTGGTGGCCCTGGCGGTGATGGCCATCTTCCTGCTCTGGTG 60
 QY 21 AspLeuMetHisArgArgGlnArgTTPAlaAlaArgTyrSerProGlyProLeu 40
 DB 61 GACCTGATGCACGGGGCCCAACGCTGGGTGACGCTACCCACGAGCCCTTGGCCTG 120
 QY 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
 DB 121 CCCGGGCTGGGAACCTGCTGATGGACTTCCAGAACACACACCATACTGCTTCGACCAG 180
 QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTTPThrProValVal 80
 DB 181 TTGGCGCGCGCTTTCGGGACGTGTTCAGCTGACGTGGCTGGACCGCGGTGGTGGT 240
 QY 81 LeuAsnGlyLeuAlaAlaValArgGlnAlaLeuValThrHisGlyGluAspThrAlaAsp 100
 DB 241 CTCATGGGCTGGCGCGCTGGCGAGGGCTGGTGACCCACGCGAGGACACCCCGGAC 300
 QY 101 ArgProProValProIleThrGlnLeuLeuGlyPheGlyProArgSerGln----- 117

Db 301 CGCCGCCCTGTGCCCATCACCAGATCTCTGGGTTTTCGGGCCCGGTTTCCCAAGGGGTGTTTC 360
 QY 117 ----- 117
 Db 361 CTGGCGCGCTATGGGCGCGCTGGCGGAGCAGAGCGGCTTCTCCGTCTCCACCTTGGCGC 420
 QY 117 ----- 117
 Db 421 AACTTTGGGCTGGGCAAGAAGTCTGTGGAGCAGTGGGTGACCGAGGAGGCGCGCTTCCTT 480
 QY 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
 Db 481 TGTGGCGGCTTCGCGCAACACCTCCCGAGCGCCCTTTTCGCCCAACGGTCTCTTGGACAA 540
 QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149
 Db 541 GCGGTGAGCAGCTGATCGCTTCCCTACCTGCGGGCGCGCTTTCGAATACGACGACCTT 600
 QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluLysSerGlyPheLeu 169
 Db 601 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGAGTGAAGGAGGAGTGGGCTTCTTCTG 660
 QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
 Db 661 CGCGAGGTGCTGAATGCTGCTCCCGCTCTCTGTCATATCCAGCGCTGGCTGGCAAGTTC 720
 QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
 Db 721 CTACGCTTCCAAAGGCTTCTGACCCAGCTGGATGAGCTGCTTAACGTGAGCAGGATG 780
 QY 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
 Db 781 ACCTGGGACCCAGCCCGAGACCTGACTGAGGCTTCTTCTGCGAGATGGAG 840
 QY 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
 Db 841 AAGCCCAAGGGGACCCCTGAGAGCAGCTTCAATGATGAGAACCTGTGCATAGTGGTGGCT 900
 QY 250 AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTTPGlyLeuLeuLeu 269
 Db 901 GACCTGTTCTTGGCGGGATGGTGACCTGACACCGCTGGCTGGGCGCTTCTGCTC 960
 QY 270 MetIleLeuHisProAspValGlnArgValGlnGlnGluIleAspAspValIleGly 289
 Db 961 ATGATCTTACATCCGATGTCGAGCGCGTGTCCAAAGGAGATCGACGACGTGATGG 1020
 QY 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
 Db 1021 CAGGTGCGGCGACAGAGATGGGTGACGAGGCTCATGCTCCCTACACCATGCGCGTATT 1080
 QY 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
 Db 1081 CATGAGGTGACGCGCTTTGGGAGCATGTCCTCCCTGGGTGTGACCATATGATCCCGT 1140
 QY 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
 Db 1141 GACATCGAAGTACAGGCTTCCGATCCTTAAGGGAACGACACTCATCACCACTGTCA 1200
 QY 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
 Db 1201 TCGGTGCTGAAGGATGAGCGCTCTGGGAGAAGCGCTTCCGCTTCCACCCCGGACACTTC 1260
 QY 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
 Db 1261 CTGGATGCCCGAGGCGCACTTTGTGAAGCGGAGGCGCTTCTCTGCTTCTTCAGAGCGCGC 1320
 QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuLeuPhePheThrSerLeu 409
 Db 1321 CGTGATGCTCGGGAGCGCTTGGCGCGCATGGAGCTTCTTCTTCTTCTTCTTCTTCTTCT 1380
 QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
 Db 1381 CTGCAGCACTTTCAGCTTCTCGGTGCCACTGGACAGCCCGCGCCCGGACCATGCTGTC 1440

Qy 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTTCTGCTGACCCCATCCCCCTATGAGCTTTGTGCTGTGCCCCGC 1491

Search completed: February 29, 2004, 00:10:25
Job time : 418 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 28, 2004, 16:51:14 ; Search time 3605 Seconds
(without alignments)
5362.273 Million cell updates/sec

Title: US-09-820-788a-2
Perfect score: 2330
Sequence: 1 MGLEALVPLAIVIAIFLLV.....HGVFAFLVTPSPYELCAVPR 446

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09820788/runat_24022004_141407_9084/app_query.fasta_1.583
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09820788 @CGN 1 1 2372 @runat_24022004_141407_9084 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pin.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2289.5	98.3	1494	6	E10868 cDNA encodi
2	2289.5	98.3	1494	6	AR399336 Sequence
3	2285.5	98.1	1494	6	E10647 Human cDNA
4	2285.5	98.1	1494	6	E10869 cDNA encodi
5	2285.5	98.1	1494	6	AR399337 Sequence
6	2282.5	98.0	1567	6	AX687029 Sequence
7	2282.5	98.0	1567	9	HSDBI
8	2282.5	98.0	1567	9	HUMCYPDB1
9	2281.5	97.9	1494	6	E10867
10	2281.5	97.9	1494	6	AR399335 Sequence
11	2277.5	97.7	1494	6	E10870
12	2277.5	97.7	1494	6	AR399338 Sequence
13	2271.5	97.5	1565	6	AX687033 Sequence
14	2267.5	97.3	1568	6	AX687030 Sequence
15	2254.5	96.7	1566	6	AR084365 Sequence
16	2250.5	96.6	1545	6	E15820 cDNA encodi
17	2241.5	96.2	1568	6	A20907
18	2241.5	96.2	1568	6	AR084372 Sequence
19	2205.5	94.6	1566	6	AR084374 Sequence
20	2197.5	94.3	1585	9	AF301911 Macaca fu
21	2141.5	91.9	1571	6	AR084373 Sequence
22	2140.5	91.9	1494	9	MFU38218 Macaca fasc
23	2140.5	91.9	1569	6	AR084375 Sequence
24	2138.5	91.8	1494	9	AY082602 Callithri
25	2093.5	89.8	1593	6	AX195173 Sequence
26	2069.5	88.8	1615	9	D29822 Callithrix
27	1908.5	81.9	1501	6	AR084376 Sequence
28	1769.5	75.9	1615	4	BTCYP2DA
29	1751.5	75.2	1503	4	AB008785
30	1739.5	74.6	1852	4	SSVD325H
31	1734.5	74.4	1538	4	BTCYP2D
32	1722.5	73.9	1632	10	AB008425
33	1719.5	73.8	1195	6	AR380887
34	1719.5	73.8	1195	9	HUMP450DB1
35	1712.5	73.5	1218	9	HSP450
36	1708.5	73.3	1778	10	RNU48219
37	1708.5	73.3	2746	6	AX401787
38	1708.5	73.3	2746	6	AX827330
39	1708.5	73.2	2746	10	RNU48220
40	1704.5	73.2	1609	10	AF221525
41	1700.5	73.0	5503	9	HUMCYP2DG
42	1698.5	72.9	1741	10	BC016256
43	1697.5	72.8	1503	4	AB008784
44	1690.5	72.5	1631	4	AB004268
45	1690.5	72.5	1665	4	DOGCP2D

ALIGNMENTS

RESULT 1

E10868
 LOCUS 1494 bp RNA linear PAT 29-SEP-1997
 DEFINITION cDNA encoding human cytochrome P450.
 E10868
 ACCESSION
 VERSION E10868.1 GI:22027962
 KEYWORDS JP 1996056695-A/17
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1494)
 Yashiki, K., Sakaki, T., Yabuseaki, Y., Komai, K., Kaneko, H. and
 Nakatsuka, I.
 TITLE METHOD FOR EVALUATING SAFETY
 JOURNAL Patent: JP 1996056695-A 17 05-MAR-1996;
 SUMITOMO CHEM CO LTD
 COMMENT OS Homo sapiens (human)
 PN JP 1996056695-A/17
 PD 05-MAR-1996
 PF 15-JUL-1994 JP 1994164184
 PR 20-JUL-1993 JP 93P 201120, 30-JUL-1993 JP 93P 208279, PR
 17-JUN-1994 JP 94P 136053
 PI HAYASHI KOJI, SAKAKI TOSHIYUKI, YABUSAKI YOSHIYASU, PI KOMAI
 KOICHIRO,
 PI KANEKO HIDEO, NAKATSUKA IWA0
 PC C12Q1/02.C12M1/34.C12Q1/26;
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No; Location/Qualifiers
 FH Key
 FH source 1..1494
 FT CDS 1..1494
 FT /organism="Homo sapiens"
 FT /product="human cytochrome P450 2D6".
 FEATURES
 source 1..1494
 /organism="Homo sapiens"
 /mol_type="genomic RNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores: 2.2e-185 Length: 1494
 Pred. No.: 2289.50 Matches: 445
 Score: 2289.50
 Percent Similarity: 89.54% Conservative: 0
 Best Local Similarity: 89.54% Mismatches: 1
 Query Match: 98.26% Indels: 51
 DB: 6 Gaps: 1
 US-09-820-788A-2 (1-446) x E10868 (1-1494)
 QY 1 MetGlyLeuGluAlaLeuValProLeuAlaValAlaValAlaValAlaValAlaVal 20
 Db 1 ATGGGGCTAGACACATGTGGCCCTGGCGGTGATAGTGGCCATCTCTCTCTCTGGTG 60
 QY 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTySerProGlyProLeuProLeu 40
 Db 61 GACCTGATGCACCGCGCCCAACCGTGGCTGCACGCTACCCACACGCGCCCTGCCACTG 120
 QY 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrcysPheAspGln 60
 Db 121 CCCGGGCTGGCAACTGTGTCATGTGGACTTCACAGAACACACCATCTCTCGACCCAG 180
 QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
 Db 181 TTGCGGCGCGCTTCGGGGACGTTGTCAGCTGCAGCTGGCGCTGGACCGCGGTGGTGGT 240
 QY 81 LeuAsnGlyLeuAlaValAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
 Db 241 CTCAATGGGCTGGCGCGCGTGGCGAGGCGCTGGTGACCCACCGCGGAGGACACCGCGCAG 300
 QY 101 ArgProProValProLeuThrGlnLeuGlyPheGlyProArgSerGln----- 117
 Db 301 CGCCCGCCCTGTGGCCCATCACCCAGATCTGGGTTCGGGCGCGGTTCACAGGGGTGTTTC 360
 QY 117 ----- 117
 Db 361 CTGGCGCGCTATGGGCGCGCTGGCGCGAGCAGAGCGGCTTCTCCGTCTCCACCTTGC 420
 QY 117 ----- 117
 Db 421 AACTTTGGGCTGGGCAAGAAGTCGTGGAGCAGTGGTGACCGAGAGGCGCGCTGCCTT 480
 QY 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
 Db 481 TGTGGCGCTTCGCCCAACCACTCCCGACGCCCTTTGCCCCCAACGGTCTCTTGGACAAA 540
 QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyTrpAspPro 149
 Db 541 GCGGTGAGCAACGTGATCGCTCCCTCACCTCGCGGCGCGCTTCCAATACGACGACCT 600
 QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLeuGluSerGlyPheLeu 169
 Db 601 CGCTTCTCAGCTCTGACCTAGCTCAGGAGGAGCTGAAGAGGAGTGGGCTTTCG 660
 QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
 Db 661 CGCAGAGTGTGAATGCTGTCCTCCCTCTCTCTGCAATATCCAGCGCTGGCTGCAAGTTC 720
 QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
 Db 721 CTACGCTTCCAAAAGGCTTCTGACCCAGCTGGATGAGTCTTAATGAGCACAGGATG 780
 QY 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
 Db 781 ACCTGGGACCCAGCCCGAGCCCGAGACCTGAGCTGAGGCTTCTCTGCGAGATGGAG 840
 QY 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
 Db 841 AAGGCCAAGGGGACCTCTGAGAGCAGCTTCATGATGAGAACTGGCATAGTGGTGGCT 900
 QY 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 269
 Db 901 GACCTGCTTCTCTCGCGGATGGTGCACCTCGACCTGGCTGGCGCTCTCTGCTC 960
 QY 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspValIleGly 289
 Db 961 ATGATCTTACATCCCGATGTGAGCGCGGTGTCCAAACAGAGATCGACCGGTAGAGG 1020
 QY 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrrThrThrAlaValIle 309
 Db 1021 CAGGTGGCGCGACGAGATGGGTGACAGGCTCACATGCCCTACACCATGCGCGTATT 1080
 QY 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
 Db 1081 CATGAGTGCAGCGCTTTCGGGACATCGTCCCGCTGGGTGTGACCCATATGACATCCCGT 1140
 QY 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
 Db 1141 GACATGAAATACAGGGCTTCGCGATCCCTAAGGGGAACGACACTCATACCAACCTGTGTC 1200
 QY 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
 Db 1201 TGGTGTCTGAAGATGAGCGCTCTGGGAGAGACCTTCCGCTTCCACCCCGACACTTC 1260
 QY 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
 Db 1261 CTGGATGCCAGGGCCACTTGTGAAGCGGAGGCTTCTGCTCTTCTCAGCAGGCGCG 1320
 QY 390 ArgAlaCysLeuGlyClnProLeuAlaArgMetGluLeuPheLeuPhePheThrSerLeu 409
 Db 1321 CGTGCAATGCTCGGGGAGCCCTCGGCCCGCATGGAGCTTCTCTCTTCTTACCTCCCTG 1380
 QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 429

Db 1381 CTGAGACATTCAGCTTCTGGTGCCCACTGGACAGCCCGCCGACCCACCATGGTGC 1440
Qy 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTCTCGGTGACCCCATCCCCCTATGAGCTTTGTGCTGCCCCGC 1491

RESULT 2

AR399336 1494 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 17 from patent US 6620593.
DEFINITION AR399336
ACCESSION AR399336
VERSION AR399336.1 GI:40141144
KEYWORDS

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 1494)

AUTHORS

Hayashi,K.; Sakaki,T., Yabusaki,Y., Komai,K., Kaneko,H. and

TITLE

Method for safety evaluation of chemical compound using recombinant

JOURNAL

yeast expressing human cytochrome P450

FEATURES

Patent: US 6620593-A 17 16-SEP-2003;

source

Location/Qualifiers

1..1494

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 2,2e-185 Length: 1494

Score: 2289.50 Matches: 445

Percent Similarity: 89.54% Conservative: 0

Best Local Similarity: 89.54% Mismatches: 1

Query Match: 98.26% Indels: 51

DB: 6 Gaps: 1

US-09-820-788A-2 (1-446) x AR399336 (1-1494)

Qy 1 MetGlyLeuAlaLeuValProLeuAlaValAlaValAlaValAlaValAlaVal 20

Db 1 ATGGGGCTAGAGCACTGGTGCCCTGGCGGTGATGATGGCCATCTCTCTCTCTGGTG 60

Qy 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40

Db 61 GACCTGATGACACCGCGCCCAACGCTGGCTGGACGCTACCCACGAGCCCTGGCCACTG 120

Qy 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60

Db 121 CCGGGCTGGCAACCTGCTGCAATGGAGCTTCCAGAACACACCATCTGCTTCGACCAG 180

Qy 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaValThrProValValVal 80

Db 181 TTGGCGCGCGCTTCGGGAGCGTTTCAGCTGACCTGGCTGGAGCGCGGTGGTGGTG 240

Qy 81 LeuAsnGlyLeuAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100

Db 241 CTCAATGGGCTGGCGCGCGCTGGCGAGCGCTGGTGAACCCACCGCGAGGACACCGCCGAC 300

Qy 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln 117

Db 301 CGCCCGCGCTGTGCCATCACCAGATCTCTGGGTTTCGGGCGCGCTTCCCAAGGGGTGTTTC 360

Qy 117 ----- 117

Db 361 CTGGCGCGCTATGGGCGCGCTGGCGCGAGCAGAGCGCTTCTCGCTCCACCTTGGCGC 420

Qy 117 ----- 117

Db 421 AACTTGGGCTGGGCAAGAAGTCGCTGGAGCAGTGGTGACCGAGGAGCGCGCTTCCTT 480

Qy 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspTys 129

|||||

Db 481 TGTGCCCGCTTCCCAACCACTCCGAGCGCCCTTTTGGCCCCAACCGTCTCTTTGGACAA 540
Qy 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149
Db 541 GCGGTGAGCAACGCTGATCGCTTCCCTACCTGGGGCGCGCTTCGNATACGACGACCT 600
Qy 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
Db 601 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGAGCTGAAGGAGGAGTGGGGCTTCTG 660
Qy 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
Db 661 CGGAGGTGCTGAATGCTGTCCCGCTCTCTCATATCCACGCGTGGCTGCAAGGTC 720
Qy 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
Db 721 CTACGGTTCCAAAGGCTTCTTGACCGAGCTGGATGAGCTGCTAATCTGAGCACAGATG 780
Qy 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 781 ACCTGGGACCCAGCCCGAGCCCTGACTGAGGCGCTTCTTGGCAGAGATGGAG 840
Qy 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValAla 249
Db 841 AAGGCCAAGGGGAACCTCGAGAGCAGCTTCAATGATGAGAACCCTGGCGCATAGTGGT 900
Qy 250 AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaValThrGlyLeuLeu 269
Db 901 GACCTGTTCTTGGCGGATGGTGACCTCGACCACTGGCGCTGGGCGCTTCTTGCTC 960
Qy 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluLeuAspValIleGly 289
Db 961 ATGATCTTACATCCGATGTCAGCGCGCTGTCCAAACAGAGAGATCCAGCGGTAGGG 1020
Qy 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
Db 1021 CAGGTGCGGCGCACAGAGATGGGTGACAGGCTCATATGCCCTTACACCATGCGGTG 1080
Qy 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATGAGTGCAGCGCTTGGGAGCATCGTCCCTTGGGTGGTGGACCATATGATATCCCGT 1140
Qy 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAGTACAGGCTTCCGATCCCTAAGGAGACGACACTCATACCAACCTGTCA 1200
Qy 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAAGGATGAGCGCTCTGGGAGAGCGCTTCCGCTTCCACCCGAACTTC 1260
Qy 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGCGCATTTGTGAAGCGGAGGCGCTTCTGCTTCTTTCAGAGCGCCG 1320
Qy 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrSerLeu 409
Db 1321 CGTGATGCTCGGGGAGCGCTTGGCGCGCATGGAGCTTCTCTCTTCTTCACTCCCTG 1380
Qy 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
Db 1381 CTGACGACCTTTCAGCTTCTCGGTGCGCATGAGACGCGCGCGCCACCATGATGTGTC 1440
Qy 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTCTTCTGCTGACCCCATCCCTTATGAGCTTTGTGCTGTCGCCGC 1491

RESULT 3

E10647

LOCUS

Human cDNA encoding cytochrome P4502B6.

DEFINITION

E10647

ACCESSION

GI:22027703

VERSION

JP 199602199-A/8.

E10647 1494 bp RNA linear PAT 29-SEP-1997

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1494)
AUTHORS Funae, Y., Imaoka, S., Matsuaki, Y., Hayashi, K. and Yabusaki, Y.
TITLE ANTIBODY RECOGNIZING CYTOCHROME P4502D6 ORIGINATED FROM MAN
JOURNAL Patent: JP 1996027199-A 8 30-JAN-1996;
SUMITOMO CHEM CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1996027199-A/8
PD 30-JAN-1996
PE 15-JUL-1994 JP 1994164186
PI FUNAE YOSHIHIKO, IMAOKA SUSUMU, MATSUKI YASUSHI, HAYASHI KOJI,
FY YABUSAKI YOSHIYASU
PC C07K16/18, C12N15/09//C12N1/19, C12N9/02, G01N33/53, G01N33/53; CC
strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1. 1494
FT organism="Homo sapiens"
FT /tissue type="liver"
FT CDS 1. 1494
FT /product="cytochrome P4502D6".
FEATURES
source Location/Qualifiers
1. 1494
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4,82e-185 Length: 1494
Score: 2285.50 Matches: 444
Percent Similarity: 89.54% Conservative: 1
Best Local Similarity: 89.34% Mismatches: 1
Query Match: 98.09% Indels: 51
DB: 6 Gaps: 1
US-09-820-788A-2 (1-446) x E10647 (1-1494)
QY 1 MetGlyLeuAlaLeuValProLeuAlaValAlaValAlaValAlaValAlaValAlaVal 20
Db 1 ATGGGCTAGAGCACTGGTGGCCCTGGCGGATAGTGGCCATCTCTCTCTCTCTCTCT 60
QY 21 AspLeuMetHisArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40
Db 61 GACCTGATGACCGCGCGCAACGCTGGCTGCACGCTACCCACGAGCCCTGCGCACTG 120
QY 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 121 CCGGGCTGGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
Db 181 TTGCGCGCGCGCTTCGGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 81 LeuAsnGlyLeuAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTCAATGGCTGGCGCGCTGGCGGAGCGCTGGTGCACCCACGCGGAGGACACCGCG 300
QY 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln----- 117
Db 301 CGCGCGCGCTGTGCCATCACCAGATCCTGGGTTTCGGGCGCGGTTCCCAAGGGGTGTT 360
QY 117 ----- 117
Db 361 CTGGCGCGCTATGGGCGCGCTGGCGGAGAGAGAGCGGCTTCTCCGCTCTCCACCTTGG 420
QY 117 ----- 117
Db 421 AACTTGGGCTGGGCAAGAGTCTGCTGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 480

QY 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
Db 481 TGTGGCGGCTTCGCAACACATCCCGAGCGCCCTTCGCCCCCAACGGTCTCTTGGACAAA 540
QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149
Db 541 GCGGTGAGCAACGTGATCGCTCCCTACCTGCGGGGCGCGCTTCGATATACGACGACCT 600
QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
Db 601 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGAGCTGAAGAGGAGTCCGGCTTCTG 660
QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
Db 661 CGCAGGCTGCTGAATGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
Db 721 CTACGCTTCCAAAGGCTTCTGACCCAGCTGGATGAGCTGCTAACTGAGCAGGATG 780
QY 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 781 ACCTGGGACCCAGCCAGCCCGGAGACCTGACTGAGGCTTCTGCTGCGCAGATGGAG 840
QY 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
Db 841 AAGGCCAAGGGGAACCTCGAGAGCAGCTTCAATGATGAGAACTCGCATGATGGTGGCT 900
QY 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 269
Db 901 GACCTGTTCTGCGGGATGGTGACCACTCGACCACTGCGCTGGCGCTGCGGCTCTCT 960
QY 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGlnIleAspAspValIleGly 289
Db 961 ATGATCTACATCCGAGTGTGAGCGCGCTGTGCAACAGAGATCGACGCTGATAGGG 1020
QY 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
Db 1021 CAGGTGCGGCGACAGATGGTGACCACTGACCACTGACCACTGACCACTGACCACTG 1080
QY 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATGAGTGCAGCGCTTGGGAGCATGTCCTCCCTGGGTGGGTGACCATATGATCCCT 1140
QY 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAAGTACAGGCTTCCGCATCCCTAAGAGGAACGACACTCATCAACCACTGTCA 1200
QY 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAAGGATGAGCGCTGTGGGAGAGCGCTTCCGCTTCCACCCGAGCACTTC 1260
QY 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGCGCCACTTGTGAAGCGGAGGCTTCTGCTGCTTCTCAGCAGCGCG 1320
QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuLeuPheLeuPheThrSerLeu 409
Db 1321 CGTGATGCTCGGGGAGCCCTGCGCGCATGAGCTTCTGCTTCTTCTTCTTCTTCTTCT 1380
QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
Db 1381 CTGAGCACTTCACTTCTCGGTGCCACTGGAGAGCGCGCGCGCGCGCGCGCGCGCG 1440
QY 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTCTGCTGAGCCCATCCCTATGAGCTTGTGCTGTGCTGCTGCTGCTGCTGCT 1491

RESULT 4
E10869
LOCUS
DEFINITION cdna encoding human cytochrome P450.

E10869 1494 bp RNA linear PAT 29-SEP-1997

ACCESSION E10869
VERSION E10869.1 GI:22027963
KEYWORDS JP 1996056695-A/18.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1494)
AUTHORS Hayashi,K., Sakaki,T., Yabusaki,Y., Komai,K., Kaneko,H. and Nakatsuka,I.
TITLE METHOD FOR EVALUATING SAFETY
JOURNAL Patent: JP 1996056695-A 18 05-MAR-1996;
COMMENT SUMITOMO CHEM CO LTD
OS Homo sapiens (human)
PN JP 1996056695-A/18
PD 05-MAR-1996
PF 15-JUL-1994 JP 1994164184
PR 20-JUL-1993 JP 93P 201120, 30-JUL-1993 JP 93P 208279, PR 17-JUN-1994 JP 94P 136053
PI HAYASHI KOJI, SAKAKI TOSHIYUKI, YABUSAKI YOSHIYASU, PI KOMAI KOICHIRO,
PI KANEKO HIDEO, NAKATSUKA IWAQ
PC C1201/02.C12M1/34.C12Q1/26;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..1494
FT CDS /organism="Homo sapiens"
FT /product="human cytochrome P450 2D6".
FEATURES
source
1..1494
/organism="Homo sapiens"
/mol_type="genomic RNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.82e-185 Length: 1494
Score: 2285.50 Matches: 444
Percent Similarity: 89.54% Conservative: 1
Best Local Similarity: 89.34% Mismatches: 1
Query Match: 98.09% Indels: 51
DB: 6 Gaps: 1
US-09-820-788A-2 (1-446) x E10869 (1-1494)
Qy 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
Db 1 ATGGGGCTAGAGACACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCTGCTCTGGTG 60
Qy 21 AspLeuMetHisArgArgGlnArgTTPAlaAlaArgTyrSerProGlyProLeuProLeu 40
Db 61 GACCTGATGACACCGCGCAACGCTGGCTGACGCTACCCACCGAGCCCTGCCACTG 120
Qy 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 121 CCGGGGCTGGGCAACCTGGCTGATGTGACTTCAGAACACACCATACTGCTTCGACCAG 180
Qy 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTTPThrProValValVal 80
Db 181 TTGGGGCGCGCTTCGGGGACGTGTTTCAGCTGCAGCTGGCCCTGGAGCCCGGTGGTG 240
Qy 81 LeuAsnGlyLeuAlaValAlaArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTCAATGGCTGGCGCGCTGGCGAGCGCTGGTGACCCACCGCGGAGGACACCGCGGAC 300
Qy 101 ArgProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln----- 117
Db 301 CGCCCGCTGTGCCCATCACCCAGATCTCGGGTTTCGGGCGCGGTTTCCCAAGGGGTGTC 360

Qy 117 ----- 117
Db 361 CTGGCGGCTATGGGCGCGCTGGCGGAGCAGAGGGCGCTTCTCCGCTCCACCTTGCGC 420
Qy 117 ----- 117
Db 421 AACTTGGGCTGGGCAAGAAAGTCCTGGAGCAGCTGGGTGACCGAGGAGCGCGCTTGCTT 480
Qy 118 ----- GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
Db 481 TGTGGCCGCTTCGCCAACCACTCGGAGCGCCCTTTCGCCCAACGCTCTCTTGGACAAA 540
Qy 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149
Db 541 GCCGTGAGCAACGTGATCGCTCCCTCACCTCGGGCGCGCTTCGAAATACGACGACCT 600
Qy 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
Db 601 CGCTTCTCAGGCTGCTGGACCTAGCTCAGAGGGGACTGAAGGAGGAGTGGGGCTTTCTG 660
Qy 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
Db 661 CGGAGGTGCTGAATGCTGTCCCGCTCTCTGCATATCCAGCGCTGGCTGGCAAGGTC 720
Qy 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
Db 721 CTACGCTTCCAAAAGGCTTTCCTGACCCAGCTGGATGAGCTGCTAACTGAGCAGCATG 780
Qy 210 ThrTTPAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 781 ACCTGGGACCCAGCCAGCCCGAGACCTGACTGAGCGCTTCTCTGGCAGAGATGGAG 840
Qy 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
Db 841 AAGCCCAAGGGGAACCTTGAGACAGCTTCAATGATGAGAACTTGGCCATAGTGGTGCT 900
Qy 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTTPGlyLeuLeuLeu 269
Db 901 GACCTGTCTCTCGCGGATGGTGACCACTCGACCACTGGCTGGCGCTCCCTGCTC 960
Qy 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspAspValIleGly 289
Db 961 ATGATCTTACATCCGGATGTGACGCGCGTGTCCAACAGAGAGATCGACGACGTATAGG 1020
Qy 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
Db 1021 CAGGTGGCGGACCCAGAGATGGGTGACAGGCTCACATGCCCCCTACACCACTGGCGT 1080
Qy 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATGAGGTGCAGCGCTTGGGACATCGTCCCTGGGTGTGACCATATGACATCCCGT 1140
Qy 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAAGTACAGGGCTTCGCACTCCCTAAGGGAACGACACTCATACCAACCTGCA 1200
Qy 350 SerValLeuLysAspGluAlaValTTPGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTCTGAAGGATCAGGCGCTCGGGAGAGCCCTTCGCTTCCACCCCGAACACTTC 1260
Qy 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGGCCACTTTGTGAAGCCGGAGGCCCTTCTGCTTCTTCAGCAGGCCGC 1320
Qy 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 409
Db 1321 CGTGATGCTCGGGGAGCCCTGGCGCGCATGGAGCTTCTCTCTTCTTCACTCCCTG 1380
Qy 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 429
Db 1381 CTGCAGCACTTCACTTCTCGTGCCCACTGGGACAGCCCGCGCCGACCATGATGGTGC 1440

REFERENCE 1 Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Milos, P.M. and Webb, S.M.

TITLE Variants of the human cyp2d6 gene

JOURNAL Patient: BP 1281755-A 3 05-FEB-2003; Pfizer Products Inc. (US)

FEATURES Location/Qualifiers

source 1..1567

ORIGIN /organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

Alignment Scores:

Pred. No.: 9,23e-185 Length: 1567

Score: 2282.50 Matches: 443

Percent Similarity: 89.54% Conservative: 2

Best Local Similarity: 89.13% Mismatches: 1

Query Match: 97.96% Indels: 51

DB: 6 Gaps: 1

US-09-820-788a-2 (1-446) x AX687029 (1-1567)

Qy 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20

Db 1 ATGGGGCTAGAGCACTGGTCCCTGGCCGTGGATGAGTGGCCATCTTCCTGCTCCTGGTG 60

Qy 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40

Db 61 GACCTGATGACCGGGCCACGCTGGCTGCAGCTACCCACAGGCCCTCCGCCACTG 120

Qy 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60

Db 121 CCCGGCTGGGCAACCTGCTCATGTGGACTCCAGAACACACATACCTGCTTCGACCAG 180

Qy 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80

Db 181 TTGGCGGCGCCCTTCGGGGAGCGTGTACGCTGAGCTGGCTGGCCGTGGAGCGCGTGTG 240

Qy 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100

Db 241 CTCATGGGCTGGGGCGCTGGCGAGCGCTGTGACCCACGCGGAGGACACCCGCCGAC 300

Qy 101 ArgProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln----- 117

Db 301 CGCCCGCTGTGCCCATCACCCAGATCTGGGTTTGGGGCGCGTTCCTCCAGGGGTGTC 360

Qy 117 ----- 117

Db 361 CTGGCGGCTATGGGCGCGCGTGGCGAGCAGAGCGCTTCTCCGTGTCCACCTTGGCG 420

Qy 117 ----- 117

Db 421 AACTTGGGCTGGGCAAGAGTCTCTGGAGCAGTGGGTGACCGAGGAGCGCGCTGCTT 480

Qy 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129

Db 481 TGTGCGGCTTCGGCAACCACTCCGGAGCGCCCTTCGCCCAACGGTCTCTTGACAAA 540

Qy 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149

Db 541 GCCGTGAGCAACGATGATCGCTCCCTCACCTGGGGCGCGCTTCGAGTACGACGACCT 600

Qy 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169

Db 601 CGCTTCTCAGCTGTGGACTAGCTCAGAGGAGTGAAGAGGAGTGGGGCTTTCG 660

Qy 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189

Db 661 CGCGAGGTGCTGAATGTGTCCCGCTCTCTGATATCCAGCGCTGGCTGGCAAGGTC 720

Qy 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209

Db 721 CTACGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGCTGCTAACTGAGCACAGGATG 780

Qy 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229

Db 781 ACCTGGGACCCAGCCAGCCCTCCCGAGACCTGACTGAGGCTTCTCTGGCAGAGATGG 840

Qy 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249

Db 841 AAGGCCAAGGGGAACCCCTGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGTGGCT 900

Qy 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeu 269

Db 901 GACCTGTCTCTGCGGGATGGTACCCACTCGACCGCTGGCTGGGGCTCTCTGCTC 960

Qy 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspAspValIleGly 289

Db 961 ATGATCTTACATCCGGATGTGCAGCGCTGTCCACAGGAGATCGACAGCTGATAGG 1020

Qy 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309

Db 1021 CAGGTGGCGGACCCAGAGATGGGTGACCCAGGCTCACATGCTCCCTACACCTGCGGT 1080

Qy 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329

Db 1081 CATGAGTGGCAGCGCTTTGGGGACATCTGCCCTGGGTATGACCCCATATGACATCCCG 1140

Qy 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349

Db 1141 GACATCGAGTACAGGCTTCCGATCCCTAAGGGAACGACACTCATCACCACCTGTCA 1200

Qy 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369

Db 1201 TCGGTGTGAAGGATGAGGCGCTCTGGGAGAAGCCCTTCGCTTCCACCCGACACTTC 1260

Qy 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389

Db 1261 CTGGATCCCGAGGCGCACTTTGTGAAGCCGAGGCTTCTCTGCTCTTCTCAGACGGCCG 1320

Qy 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 409

Db 1321 CGTGCATGCTTCGGGAGCCCTTGGCCGATGAGCTCTTCTCTTTCACCTCCCTG 1380

Qy 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 429

Db 1381 CTGCAGCACTTACGCTTCTCGGTGCCACTGGACAGCCCGCCGCCACCATGCTGTC 1440

Qy 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446

Db 1441 TTTGCTTCTGTTGAGCCCATCCCTATGAGCTTTGTGCTGTGCGCCGC 1491

RESULT 7

LOCUS HSDB1 1567 bp mRNA linear PRI 15-JUN-2000

DEFINITION Homo sapiens mRNA for cytochrome P450 db1.

ACCESSION X08006.Y00300

VERSION X08006.1 GI:30450

KEYWORDS cytochrome P450; cytochrome P450 db1.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1567)

AUTHORS Gonzalez, F.J., Skoda, R.C., Kimura, S., Umeno, M., Zanger, U.M., Nebert, D.W., Gelboin, H.V., Hardwick, J.P. and Meyer, U.A.

TITLE Characterization of the common genetic defect in humans deficient in debrisoquine metabolism

JOURNAL Nature 331 (6155), 442-446 (1988)

MEDLINE 88122614

PUBMED 3123997

REFERENCE 2

AUTHORS Nebert, D.W.

TITLE Direct Submission

JOURNAL Submitted (30-JUN-1988) Nebert D.W., NIH, Bethesda, Maryland 20892

COMMENT see also X07618 (variant a), X07619 (variant b) and X07620 (variant b').

FEATURES

source Location/Qualifiers
1. .1567
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="22"
/map="22q11.2-qter"
/cell_type="hepatocyte"
/tissue_type="liver"
/clone_lib="lambda gtl1"
1. .1494
/codon_start=1
/product="cytochrome P450 db1"
/protein_id="CAA30807.1"
/db_xref="GI:30451"
/db_xref="GOA:P10635"
/db_xref="SWISS-PROT:P10635"
/translation="MGLEALVPLAVIAVAILLELLVLMHRRQWAAAYPPGPLEPLGLG
NLHVDFTQNTFYCFDQRRRFVDFSLQIAWTPVVLNGLAAVREALVTHGEDTADRP
PVEITILGFPSPNGVFLARYGPAWRQSFVSTRLNIGLGGKSLQWVTEAAACL
CAAFANHSRPFPPNGVFLARYGPAWRQSFVSTRLNIGLGGKSLQWVTEAAACL
FLREVLNAPVLIHI PALAKVLRFOKAFITOLDLLEHMTWDPQPRDITAEFL
AEMEKAGNPESFNDENLRI VVADLFSGMVTTSITLAWGLLMLILPDVORRVOE
IDRVIGQVRPENGQDQHPYTTAVIHEVORFGDIVPLGWTMTSRDIEVQGRIPKG
TEUTLNLSVLKDEAWERKFRPHPEHFLDAQHFVKPEAFLPFSAGRACLGELPLAR
MELFLFTSLILHFSFVPTQPRPHHGFVFLVSPVYELCAVR"
1546. .1550
/notes="putative"
1567
/note="polyA site"

CDS

polyA_signal

polyA_site

ORIGIN

Alignment Scores:

Pred. No.: 9,23e-185 Length: 1567
Score: 2282.50 Matches: 443
Percent Similarity: 89.54% Conservativeness: 2
Best Local Similarity: 89.13% Mismatches: 1
Query Match: 97.96% Indels: 51
DB: 9 Gaps: 1

US-09-820-788A-2 (1-446) x HSD81 (1-1567)

Qy 1 MetGlyLeuGluAlaLeuValProLeuAlaValAlaValAlaValAlaValAlaVal 20
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCCGTGATAGTGGCCATCTTCTCGCTCGTG 60
Qy 21 AspLeuMetHisArgArgGlnArgTTPAlaAlaArgTyrSerProGlyProLeu 40
Db 61 GACCTGATGCACGGCGCCCAACGCTGGCGTGCACGCTACCCACAGGCCCTCGCCACTG 120
Qy 41 ProGlyLeuGlyAsnLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 121 CCGGGCTGGCAACCTGCTGCATGTGCACTTCCAGAACACACCATACTGCTTCGACCA 180
Qy 61 LeuArgArgPheGlyAspValPheSerLeuGlnLeuAlaValThrProValValVal 80
Db 181 TTGGCGCGCGCTTCGGGGACGCTTTCAGCGCTGCAGCTGGCGCTGGAGCGCGGTGCTG 240
Qy 81 LeuAsnGlyLeuAlaValArgGluAlaValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTCNAATGGCTGGCGCGCGCTGGCGAGCGCTGTGTGACCCACGCGCAGGACACCGCCG 300
Qy 101 ArgProValProLeuThrGlnLeuGlyPheGlyProArgSerGln----- 117
Db 301 CGCCCGCTGTGCCCATCACCCAGACTCTGGTTTTCGGCGCGCTTCCCAAGGGGTGTC 360
Qy 117 ----- 117
Db 361 CTGGCGCGCTATGGCGCGCGTGGCGGAGCAGAGCGGCTTCTCCGTGCCACCTTCGCG 420
Qy 117 ----- 117

RESULT 8
HUMCPDB1

LOCUS HUMCYPDB1 1567 bp mRNA linear PRI 02-NOV-1994
DEFINITION Human cytochrome P450 db1 mRNA, complete cds.
ACCESSION M20403.1 M19697
VERSION M20403.1 GI:181349
KEYWORDS cytochrome P450; debrisoquine 4-hydroxylase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1567)
AUTHORS Gonzalez.F.J., Vilbois.F., Hardwick.J.P., McBride.O.W.,
Nebert.D.W., Gelboin.H.V. and Meyer.U.A.
TITLE Human debrisoquine 4-hydroxylase (P450IID1): cDNA and deduced amino
acid sequence and assignment of the CYP2D locus to chromosome 22
JOURNAL Genomics 2 (2), 174-179 (1988)
MEDLINE 88314109
* PUBMED 3410476
COMMENT Original source text: Human hepatocyte, cDNA to mRNA.
Draft entry and printed copy of sequence for [1] kindly provided by
D.W.Nebert, 15-JUL-1988.
FEATURES
source
1..1567
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="22q11.2-qter"
1..1567
/genes="CYP2D"
1..1567
/genes="CYP2D"
/product="CY P450 db1 mRNA"
1..1494
/notes="debrisoquine 4-hydroxylase"
/codon_start=1
/protein_id="AA52153.1"
/db_xref="GI:181350"
/db_xref="GDB:G00-119-832"
/translations="MGLEALVPLAVIVAIFFLLVLDLMHRRQWAARYPPGPIPLPGIG
NLLHVDFTPTCYCDLRRFDFVSLQAWTPVVLNGLAAVREALVTHGEDADRP
PVPITDILGPGSQSVFLARVGPAMREORFESVTLNGLGKSLQWVTEAACL
CAFANHSGRPFPPNGILLKAVSNVITSLTGRFRYDDPRFLRLDLAQGLKEBSG
FLRELVNVLVLIHPALGKRVKFORAFULTQDELLTBHMTWDPQPPRLDTEAPL
AEMEKAGPESSEFNENLRIVVADEFSAGMTVTSTLAWGLLMLTHPDVQRRVQOE
IDIVIQQVRPPEQDARHPYTVAVIHEVQFGDIVPLGTMHTMSRDIEVQGFRI PKG
TTLITLSSVLKDEAVWEKPFPHPEHFLDAQGHFKPEAFPLFSAAGRAACGLGEPLAR
MELFLFTSLQLHFSFVSPTGQPPRSHGVFAFLVSPSYELCAVR"
ORIGIN Chromosome 22.
Alignment Scores:
Pred. No.: 9,23e-185 Length: 1567
Score: 2282.50 Matches: 443
Percent Similarity: 89.54% Conservatives: 2
Best Local Similarity: 89.13% Mismatches: 1
Query Match: 97.96% Indels: 51
DB: 9 Gaps: 1
US-09-820-788A-2 (1-446) x HUMCYPDB1 (1-1567)
Qy 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
Db 1 ATGGGGCTAGAAGCACTGTGGCCCTGGCGGTGATAGTGCCCATCTTCTGCTCTCTGGTG 60
Qy 21 AspleuMethisargArgGlnArgTtpAlaAlaArgTyrSerProGlyProLeuProLeu 40
Db 61 GACCTGATGACCGGGCCCAACGCTGGGCTGCACGCTACCCACCGCCCTGGCCACTG 120
Qy 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 121 CCGGGCTGGGCAACCTGCTGATGTGGACTTCCAGNACACACATACCTGCTTGGACCAG 180
Qy 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTtpThrProValValVal 80

181 TTGCGCGCGCGCTTCTGGGAGACGTGTTACGCTGCAGCTGGCGCTGGAGCGCGGTGGTGTG 240
Qy LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
241 CTCAATGGGCTGGCGCGCGTGGCGAGGCGCTGGTGACCCACCGCGAGGACACCGCGAC 300
Qy ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln----- 117
301 CGCGCGCTGTGGCCCATACCCAGATCTCTGGGTTTCTGGGCGCGCTTCCCAAGGGGTGTTTC 360
Qy ----- 117
361 CTGGCGCGCTATGGGCGCGCTGGCGGAGAGAGCGCTTCTCCGTGTCCACCTTGGCGC 420
Qy ----- 117
421 AACTTGGGCTTGGGCAAGAAGTCGCTGGAGCAGTGGGTGACCGAGGAGCGCGCTGCCCTT 480
Qy ----- GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
481 TGTGCGCGCTTGGCCCAACCACTCCGAGCGCGCTTTCGCGCCCAACCGGTCTCTTGGACAA 540
Qy AlavalSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149
541 GCGGTGAGCAACGTGATCGCTCCCTCACCTGCGGCGCGCTTCGAGTACGACGACCGCT 600
Qy ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
601 CGCTTCTCAGGCTGCTGACCTAGCTCAGGAGGACTGAGAGGAGGTGCGGCTTCTG 660
Qy ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
661 CGCAGGTGCTGAATGCTGCTCCCGTCTCTCGATATCCAGCGCTGGGTGCAAGGTC 720
Qy LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspLeuLeuLeuThrGluHisArgMet 209
721 CTACGCTTCCAAAGGCTTCTCGACCCAGCTGGATGAGCTGCTAACTGACGACAGATG 780
Qy ThrTrpAspProAlaGlnProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
781 ACCTGGGACCCAGCCAGCCCGCCCGAGACCTGACTGAGGCTTCTCTGGCAGAGATGGAG 840
Qy LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
841 AAGGCCAAGCGGAAACCTCAGAGCAGCTTCAATGATGAGAACTCGGCATAGTGGTGGCT 900
Qy AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTtpGlyLeuLeuLeu 269
901 GACCTGTCTCTGCGCGGATGGTGACCACTCGACCGCTGGCGCTGGCGCTCTCTGCTC 960
Qy MetIleLeuHisProAspValGlnArgArgValGlnGlnGluLeuAspValIleGly 289
961 ATGATCTTACATCCGATGTGAGCGCGCTGTCACACAGAGATCGACGACGTATAGGG 1020
Qy GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
1021 CAGGTGGCGCGACAGAGATGGTGACAGGCTCATGCTCCCTACACCATGCGCGTATT 1080
Qy HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
1081 CATGAGGTGACGCGCTTGGGAGACATCGTCCCGCTGGGTATGACCCATATGACATCCCGT 1140
Qy AspileGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
1141 GACATCGAAGTACAGGGCTTCCGATCCCTAAGGGAACGACACTCATACCAACCTGTCA 1200
Qy SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
1201 TCGGTGCTGAGGATGAGCGCGCTGAGGAGAGCCCTTCCGCTTCCACCCCGAACACTTC 1260
Qy LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
1261 CTGGATGCCAGGGCCACTTTGTGAAGCGGAGGCGCTTCTGCTTCTCTCAGCAGCGCGC 1320

```
QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeu 409
Db 1321 CGTGATGCTCGGGGAGCCCTGGCCGCATGGAGCTTCTCTTCTTCACTCCCTG 1380
QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
Db 1381 CTGAGCACTTCAAGCTTCTCGGTGCCACTGGACAGCCGCCAGCCACCATTGGTGC 1440
QY 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1441 TTGTCTTCTCGTGAGCCCATCCCTATGAGCTTGTGTGCTGCCCGC 1491

RESULT 9
E10867
LOCUS E10867 1494 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding human cytochrome P450.
ACCESSION E10867
VERSION E10867.1 GI:22027961
KEYWORDS JP 1996056695-A/16.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1494)
AUTHORS Hayashi,K., Sakaki,T., Yabusaki,Y., Komai,K., Kaneko,H. and
Nakatsuoka,I.
TITLE METHOD FOR EVALUATING SAFETY
JOURNAL Patent: JP 1996056695-A 16 05-MAR-1996;
SUMITOMO CHEM CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1996056695-A/16
PD 05-MAR-1996
PF 15-JUN-1994 JP 1994164184
PR 20-JUL-1993 JP 93P 201120, 30-JUL-1993 JP 93P 208279, PR
17-JUN-1994 JP 94P 136053
PI HAYASHI KOJI, SAKAKI TOSHIYUKI, YABUSAKI YOSHIYASU, PI KOMAI
KOICHIRO,
PI KANEKO HIDEO, NAKATSUKA IWAO
PC C12Q1/02,C12M1/34,C12Q1/26;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1. .1494
FT FT /organism='Homo sapiens'
FT CDS 1. .1494
FT /product='human cytochrome P450 2D6'.
FEATURES
source
Location/Qualifiers
1. .1494
/mol_type='genomic RNA'
/db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.: 1.06e-184 Length: 1494
Score: 2281.50 Matches: 444
Percent Similarity: 89.34% Conservative: 0
Best Local Similarity: 89.34% Mismatches: 2
Query Match: 97.92% Indels: 51
DB: 6 Gaps: 1

US-09-820-788A-2 (1-446) x E10867 (1-1494)

QY 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleAlaIlePheLeuLeuVal 20
Db 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCGGTGATGGCCATCTTCTGCTCTCTGGTG 60
QY 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeu 40
Db 21 ATGGGGCTAGAGCACTGGTGGCCCTGGCGGTGATGGCCATCTTCTGCTCTCTGGTG 60
```

```
Db 61 GACCTGATGCACCGCGCCAAACGCTGGGCTGCACGCTACCCACGAGGCCCTTGCCTG 120
QY 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 121 CCCGGCTGGGCAACCTCTGCTGATGTGACTTTCAGAAACACACCATCTGCTTCGACCAG 180
QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaIleThrProValVal 80
Db 181 TTGGCGCGCGCTTCGGGGACGTGTTCAGCTGCGAGCTGGCTGGACCGCGTGGTCTGTG 240
QY 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTCAATGGGCTGGCGCGCTGGCGAGCGCTGGTGCACCCACGCGAGGAGACACCGCCGAC 300
QY 101 ArgProProValProIleThrGlnLeuGlyPheGlyProArgSerGln----- 117
Db 301 CGCCGCGCTGTGGCCCATCACCAGATCTGGGTTCGGGGCCGCTTCCAGGGGGTGTTC 360
QY 117 ----- 117
Db 361 CTGGCGCGCTATGGCGCGCTGGCGGAGCAGAGCGCTTCTCCGTCTCCACCTTGGCG 420
QY 117 ----- 117
Db 421 AACTTGGGCTGGGCAAGAGTCTGCTGGAGCAGTGGGTGACCGAGGAGCGCGCTGCGCTT 480
QY 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
Db 481 TGTGCGCGCTTGGCCCAACCACTCCGCGAGCGCCCTTTCGCCCCAACGGTCTCTTGGACAAA 540
QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGlyLeuTyrAspAspPro 149
Db 541 GCGGTGACCAACGCTGATCGCTCCCTCACCTGCGGGCGCGCTTCGATATACGACGACCT 600
QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
Db 601 CGCTTCCTCAGCGTCTGCGACTAGCTCAGGAGGAGTGAAGGAGGAGTCTGGGGCTTCTG 660
QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
Db 661 CGCGAGGTGCTGAATGCTGTCCCGCTCTCTCGCATATCCAGCGCTGGCTGGCAAGGTC 720
QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
Db 721 CTACGCTTCCAAAGGCTTCTTGACCCAGCTGGATGAGCTGCTAACTGAGCAGCAGGATG 780
QY 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 781 ACCTGGGACCCAGCCAGCCCGCGAGACCTGACTGAGGCGCTTCTGCGCAGAGATGGAG 840
QY 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValAla 249
Db 841 AAGCCCAAGGGGAAACCTTGAGAGCAGCTTCAATGATGAGAACTCTGTCATAGTGTGGCT 900
QY 250 AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTrpGlyLeuLeuLeu 269
Db 901 GACCTGTCTCTGCGGGATGGTGACCACCTCGACACGCTGGCTGGCGCTTCTGCTGCTC 960
QY 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspAspValIleGly 289
Db 961 ATGATCTTACATCCGGATGTGCGCGCGCTGTCCAAACAGGAGATCGACGACGCTGATAGG 1020
QY 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
Db 1021 CAGGTGCGGGGACCAAGATGGGTGACCGGCTCACATGCGCTCACACCTGCGCGCTGATT 1080
QY 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATGAGTGCAGCGCTTGGGGACATCTGCCCCCTGGGTGTGACCCATATGACATCCCGT 1140
QY 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAAGTACAGGGCTTCCGATCCCTAAGGGAACGACACTCATCAACCAACCTGTCA 1200
```


QY 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAAGGATGAGGCGCTGGGAGAGCCCTTCGCTTCCACCCGAACTTC 1260
QY 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGCGCACCTTGTGAAGCCGAGGAGCTTCTGCTTTCTCAGCAGGCCGC 1320
QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 409
Db 1321 CGTGATGCTCGGGAGCCCTTGGCCGATGAGCTCTTCTCTTCTTCTTCTTCTTCTTCT 1380
QY 410 LeuGlnHisPheSerPheSerValProThrGlnProArgProSerHisGlyVal 429
Db 1381 CTCGACGACTTCAGCTTCTCGGTGCCACTTGGACAGCCCGCCAGCCACCATGGTGTG 1440
QY 430 PheAlaPheLeuValThrProSerProTyTrpGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTTCTGCTGACCCCATCCCTATGAGCTTTGTGCTGTGCCCGC 1491
RESULT 10
AR399335
LOCUS AR399335 1494 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 16 from patent US 6620593.
ACCESSION AR399335
VERSION AR399335.1 GI:40141142
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1494)
AUTHORS Hayashi, K., Sakaki, T., Yabuseki, Y., Komai, K., Kaneko, H. and Nakatsuka, I.
TITLE Method for safety evaluation of chemical compound using recombinant yeast expressing human cytochrome P450
JOURNAL Patent: US 6620593-A 16 SEP-2003;
FEATURES Location/Qualifiers
1..1494
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1,06e-184 Length: 1494
Score: 2281.50 Matches: 444
Percent Similarity: 89.34% Conservative: 0
Best Local Similarity: 89.34% Mismatches: 2
Query Match: 97.92% Indels: 51
DB: 6 Gaps: 1
US-09-820-788A-2 (1-446) x AR399335 (1-1494)
QY 1 MetGlyLeuGluAlaLeuValProLeuAlaValileValAlaPheLeuLeuVal 20
Db 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCGCTGATAGTGGCCATCTTCTGCTCTCTGGTG 60
QY 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTy-SerProGlyProLeuProLeu 40
Db 61 GACCTGATGACCGCGCCCAACGCTGGCTGACGCTACCCACAGGCGCCCTTGGCCACTG 120
QY 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyTrpCysPheAspGln 60
Db 121 CCGGGCTGGGCAACTGCTGCTGATGTGGACTTCCAGACACACCATCTGCTTGGACCA 180
QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValVal 80
Db 181 TTGCGGCGCGCTTGGGGAGCGTGTTCAGCTTCAGCTGGCGCTGGAGCGCGGTGGTGTG 240
QY 81 LeuAsnGlyLeuAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTAATGGGCTGGCGCGCTGGCGAGGCGCTGGTGACCCACGCGGAGGACACCGCGCGAC 300

QY 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln 117
Db 301 CGCCGCGCTGTGCCCATCACCAGATCTCTGGGTTCCTGGGCCGCTTCCCAAGGGGTGTTTC 360
QY 117 117
Db 361 CTGGCGCGCTATGGGCGCGCTGGCGGAGAGGCGCTTCTCCGCTCTCCACCTTGGCGC 420
QY 117 117
Db 421 AACTTTGGGCGCTGGCAAGAGTCGCTGGAGCAGTGGGTGACCGAGGAGCGCGCTGCCTT 480
QY 118 118
Db 481 TGTGGCGCTTCCGCCAACACCTCCGAGACGCCCTTTTCGCCCAACAGCTCTCTTGGACAAA 540
QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTrpAspPro 149
Db 541 GCCGTGAGCAACGTATCGCTCCCTTCCCTCGGGCGCGCTTCCGAAATACGACGACCT 600
QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
Db 601 CGCTTCTCAGGCTGCTGGACCTAGCTCAGGAGGAGCTGAAGAGGAGTCCGGCTTCTG 660
QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
Db 661 CGCGAGGTGCTGAATGCTGTCCTCCGCTCTCTTCGCATATCCAGCGCTGGCTGGCAAGGTC 720
QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
Db 721 CTACGCTTCCAAAGAGCTTTCTGACCCAGCTGGATGAGTGTAACTAGACACAGATG 780
QY 210 ThrTpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 781 ACCTGGGACCCAGCCCGAGCCCTCCGAGACCTAGCTGAGGCGCTTCTTGGCAGAGATGGAG 840
QY 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
Db 841 AAGGCCAAGGGGAACTCTGAGAGCAGCTTCAATGATGAGAACTGTGCTAGTGTGGCT 900
QY 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 269
Db 901 GACCTGTCTCTGCGGGATGGTACCTCCGACCCAGCTGGCGCTGGCGCTCTCTGCTC 960
QY 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspValIleGly 289
Db 961 ATGATCTTACATCCGATGTGGAGCGCTGTCCACAGAGAGATCGACCGCTGATAGGG 1020
QY 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyTrpThrThrAlaValile 309
Db 1021 CAGGTGCGCGACACAGAGATGGGTGACAGGCTCATGCTCCCTACACCATGCGGTGATT 1080
QY 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATGAGGTGACGCGCTTTGGGAGCATCGTCCCGCTGGGTGTGACCCATATGACATCCCGT 1140
QY 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAAGTACAGGCTTCCGCTATCCCTAAGGGAACGACACTCATACCAACCTGTCA 1200
QY 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAAGGATGAGGCGCTTGGGAGAGAGCCCTTCCGCTTCCACCCGAACTTC 1260
QY 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGCGCACCTTGTGAAGCCGAGGCGCTTCTGCTTCTTCTTCTTCTTCTTCT 1320
QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 409
Db 1321 CCGTATGCTGCGGAGGCGCTTGGCGCGAGTGGAGCTTCTTCTTCTTCTTCTTCTTCTTCT 1380
QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429

```

Db 1381 CTGAGCACTTCAGCTTCCTCGTTCCTGCTGACCACTGGACAGCCCGCCGACCACTGATGTC 1440
Qy 430 PheAlaPheLeuValThrProSerProTyrGluLeuLeuValProArg 446
Db 1441 TTTGCTTTCTGCTGACCCCATCCCTATGAGCTTTGTGCTGTCGCCGC 1491

RESULT 11
E10870
LOCUS CDNA encoding human cytochrome P450. linear PAT 29-SEP-1997
DEFINITION E10870
ACCESSION E10870
VERSION 1 GI:22027964
KEYWORDS JP 1996056695-A/19.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE Hayashi, K., Sakaki, T., Yabusaki, Y., Komai, K., Kaneko, H. and
AUTHORS Nakatsuka, I.
TITLE METHOD FOR EVALUATING SAFETY
JOURNAL SUMITOMO CHEM CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1996056695-A/19
PD 05-MAR-1996
PF 15-JUL-1994 JP 1994164184
PR 20-JUL-1993 JP 93P 201120, 30-JUL-1993 JP 93P 208279, PR
17-JUN-1994 JP 94P 136053
PI HAYASHI KOJI, SAKAKI TOSHIYUKI, YABUSAKI YOSHIYASU, PI KOMAI
KOICHIRO.
PI KANEKO HIDEO, NAKATSUKA IWAO
PC C12Q1/02,C12M1/34,C12Q1/26;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1. .1494
FT FT /organism='Homo sapiens'
FT CDS 1. .1494
FT /product='human cytochrome P450 2D6'.
FEATURES
source
Location/Qualifiers
1. .1494
/organism='Homo sapiens'
/mol_type='genomic RNA'
/db_xref='taxon:9606'
ORIGIN
Alignment Scores:
Pred. No.: 2,31e-184 Length: 1494
Score: 2277.50 Matches: 443
Percent Similarity: 89.34% Conservative: 1
Best Local Similarity: 89.13% Mismatches: 2
Query Match: 97.75% Indels: 51
DB: 6 Gaps: 1
US-09-820-788A-2 (1-446) x E10870 (1-1494)
Qy 1 MetGluValGlnAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
Db 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCGGCTGAGTAGTGGCCATCTTCCTGCTCTGGTG 60
Qy 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40
Db 61 GACCTGATGCA CGGGGCCCAACGCTGGGCTGCAAGCTACCCAGGCCCCCTGCCACTG 120
Qy 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 121 CCGGGCTGGGAACCTGCTGATGAGCTTCCAGAACACACCATACTGCTTCGACCAG 180

```

```

Qy 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaIleThrProValValVal 80
Db 181 TTGGGGCGCCGCTTCGGGGACGTTT CAGCTCTGAGCTGGCTGGACCGCGGTGGTCTGTG 240
Qy 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTCAATGGGCTGGGGCGCTGGCGAGCGCTGGTGACCCACGCGGAGACACCGCCGAC 300
Qy 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln- 117
Db 301 CGCCGCGCTGTGCCCATCACCAGATCTGGGTTCGGGGCCGCTTCCCAAGGGGTGTTC 360
Qy 117 ----- 117
Db 361 CTGGCGCGCTATGGCGCGCTGGCGGAGCAGAGCGCTTCTCGTCTCCACCTTGGCG 420
Qy 117 ----- 117
Db 421 AACTTGGGCGCTGGGCAAGAGTCTGCTGGAGCAGTGGGTGACCGAGGAGCGCGCTGCCT 480
Qy 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
Db 481 TGTGCGCGCTTTCGCCAACCACTCCGGAGCGCCCTTTGCCCCAACGGTCTCTTGGACAA 540
Qy 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149
Db 541 GCGGTGACCAACGTGATCGCTCCCTACCTGCGGGCGCGCTTCGANTACGACGACCT 600
Qy 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluLeuSerGlyPheLeu 169
Db 601 CGTCTCTCAGGCTGCTGAGCTAGCTCAGGAGGAGTGAAGGAGGAGTGGGCTTCTG 660
Qy 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyVal 189
Db 661 CGCGAGGTGCTGAATGCTGTCTCCGCTCTCTGCAATATCCGAGCGCTGGCTGGCAAGT 720
Qy 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
Db 721 CTACGCTTCCAAAGGCTTCTGACCCAGCTGGATGAGCTGCTAACTAGACACAGGATG 780
Qy 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 781 ACCTGGGACCCAGCCCGCCCGAGACCTGACTGAGGCGCTTCTGGCAGAGATGGAG 840
Qy 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAenLeuArgIleValValala 249
Db 841 AAGCCCAAGGGGAAACCCCTGAGAGCAGCTTCAATGATGAGAACCTGTGCTAGTGGT 900
Qy 250 AspLeuPheSerAlaGlyMetValThrThrThrThrLeuAlaTrpGlyLeuLeuLeu 269
Db 901 GACCTGTTCTCTGCGGGATGGTGNACCACCTCGACCGCTGGCTGGGCTCTCTGCTC 960
Qy 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGlnIleAspAspValIleGly 289
Db 961 ATGATCTTACATCCGATGTGCGCGCTGTCCAAACAGGAGATCCAGCAGCTGATAGGG 1020
Qy 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
Db 1021 CAGGTGGCGGACCCAGAGATGGGTGACAGGCTCACATGCCCTACACCACTGCCGTGATT 1080
Qy 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATGAGGTGCAGCGCTTTGGGACATCTGTCCTCCCTGGGTGTGACCCATATGACATCCGT 1140
Qy 330 AsplleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAAGTACAGGCTTCCGATCCTTAAGGGAACGACACTCATCAACACCTGTCA 1200
Qy 350 SerValLeuLysAspGluAlaValTrpGlnLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAAGGATGAGGCGCTCGGGAGAGCGCTTCCGCTTCCACCCCGAAACACTTC 1260
Qy 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389

```

```
Db 1261 CTGGATGCCAGGCGCCATTGTGAAGCCGGAGGCGCTTCTGCTTTCTCAGCAGCGCCG 1320
QY 390 ArgAlaCysLeuGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 409
Db 1321 CGTGATCCCTCGGGAGCCCTTGGCCGATGAGAGCTTCTCTCTTCTTCACTCCCTG 1380
QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 429
Db 1381 CTGCAGCACTTCAGCTTCTCGGTGCCCACTGGACAGCCCGGCCAGCCACCATGGTGC 1440
QY 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTTCTGTGGAGCCCATCCCCCTATGAGCTTTGTGCTGTGCCCGC 1491

RESULT 12
AR399338
LOCUS AR399338 1494 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 19 from patent US 6620593.
ACCESSION AR399338
VERSION AR399338.1 GI:40141148
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE 1 (bases 1 to 1494)
AUTHORS Hayashi, K., Sakaki, T., Yabuseaki, Y., Komai, K., Kaneko, H. and Nakatsuka, I.
TITLE Method for safety evaluation of chemical compound using recombinant yeast expressing human cytochrome P450
JOURNAL Patent: US 6620593-A 19 16-SEP-2003;
FEATURES Location/Qualifiers
source 1..1494
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 2,31e-184 Length: 1494
Score: 2277.50 Matches: 443
Percent Similarity: 89.34% Conservative: 1
Best Local Similarity: 89.13% Mismatches: 2
Query Match: 97.75% Indels: 51
DB: 6 Gaps: 1

US-09-820-788A-2 (1-446) x AR399338 (1-1494)
QY 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
Db 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCCGTGATAGTGGCCATCTTCTGCTCTCTGGTG 60
QY 21 AspLeuMetHisArgArgGlnArgTtpAlaAlaArgTyrSerProGlyProLeuProLeu 40
Db 61 GACTGATGACCGGGCCCAACGCTGGCTGACGCTACCCACGAGCCCTTGCACCTG 120
QY 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 121 CCCGGGCTGGCAACCTGCTGCATGTGGACTTCCAGAACACACACATACCTGCTTCGACCAG 180
QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTtpThrProValValVal 80
Db 181 TTGCGGCGCGCTTCGGGGAGCGTGTTCAGCCTGCAGCTGGCCCTGGAGCGCGGTGGTCTG 240
QY 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTCAATGGGCTGGCGCGCGTGGCGAGCGCGCTGGTGAACCCAGCGGAGACACCGCGGAC 300
QY 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln----- 117
Db 301 CGCGCGCTGTGGCCATCACCGAGATCCTGGGTTCGGGCGCGGTTCGCCAGGGGTGTC 360
QY 117 ----- 117
```

```
Db 361 CTGGCGCGCTATATGGGCCCCGCGTGGCGCGAGCAGAGGCGCTTCTCCGCTCTCCACCTTGGCG 420
QY 117 ----- 117
Db 421 AACTTTGGGCTTGGCGAAGAGTCGCTGGAGCAGTGGGTGACCGAGGAGCGCGCTTGCCTT 480
QY 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
Db 481 TGTGCGCGCTTGGCCAAACCACTCCGGAGCGCCCTTTGCGCCCAACCGGTCTCTTGGACAA 540
QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluThrAspAspPro 149
Db 541 GCGTGAGCAACGTATCGCTCCCTCCCTCACCTGCGGGCGCGCTTCCGAATACGACGACCT 600
QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluLeuSerGlyPheLeu 169
Db 601 CGCTTCTCAGGCTGCTGGACCTAGCTCAGGAGGAGTGAAGGAGGAGTGGGCTTCTG 660
QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
Db 661 CGCGAGGTGCTGAATGCTGCTCCCGCTCTCTGCATATCCAGCGCTGGCTGGCAAGTGC 720
QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
Db 721 CTACGCTTCCAAAGGCTTCTGACCCAGCTGGATGAGTGTAACTGAGCACAGATG 780
QY 210 ThrTtpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 781 ACCTGGGACCCAGCCCGAGCCCTTCTGAGGAGCTGCTGAGGCTTCTTGGCAGAGATGGAG 840
QY 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
Db 841 AAGGCCAAGGGAAACCTCAGAGCAGCTTCAATGATGAGAACTGTGTGCTAGTGGTGGCT 900
QY 250 AspLeuPheSerAlaGlyMetValThrThrSerThrLeuAlaTtpGlyLeuLeuLeu 269
Db 901 GACCTGTTCTTGGCGGATGGTACCCTGACCCAGCTGGCGCTGGGCGCTTCTTGGCTC 960
QY 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspAspValIleGly 289
Db 961 ATGATCTACATCCGATGTGAGCGCGCTGTCCACACAGAGATCCAGCAGCTGATAGGG 1020
QY 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
Db 1021 CAGGTGCGGCGCACAGAGATGGTGCACAGGCTCATCATGCCCTACACCATGCGCGTATT 1080
QY 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATAGGTGCAGCGCTTTGGGGACATCGTCCCCCTGGGTGTGACCATATGACATCCCCG 1140
QY 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
Db 1141 GACATCGAAGTACAGGGCTTCCGCATCCCTAAGGGNACACACTCATACCACTGTGCA 1200
QY 350 SerValLeuLysAspGluAlaValTtpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAAGATGAGGCGCTTGGGAGAGCGCTTCCGCTTCCACCCGAACTTC 1260
QY 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGGCCACTTTGTGAAGCGCGAGGCGCTTCTGCTTTCTCAGCAGCGCGC 1320
QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 409
Db 1321 CGTGCAATGCTCGGGAGCGCCCTGGCCCGCATGGAGCTTCTCTCTTCTTCACTCCCTG 1380
QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyVal 429
Db 1381 CTGCAGCACTTCACTTCTCGGTGCCCACTGGAGACCGCCCGGCCAGCCACCATGGTGC 1440
QY 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTTCTGTGAGCCCATCCCCCTATGAGCTTTGTGCTGTGCCCGC 1491
```

```
RESULT 13
AX687033
LOCUS AX687033 1565 bp DNA linear PAT 31-MAR-2003
DEFINITION Sequence 7 from Patent EP1281755.
ACCESSION AX687033
VERSION AX687033.1 GI:29409535
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Milos, P.M. and Webb, S.M.
TITLE Variants of the human cyp2d6 gene
JOURNAL Patent: EP 1281755-A 7 05-FEB-2003;
Pfizer Products Inc. (US)
FEATURES
source Location/Qualifiers
1..1565
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 7,96e-184 Length: 1565
Score: 2271.50 Matches: 443
Percent Similarity: 89.36% Conservative: 2
Best Local Similarity: 88.96% Mismatches: 1
Query Match: 97.49% Indels: 52
DB: Gaps: 1
US-09-820-788A-2 (1-446) x AX687033 (1-1565)
QY 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
Db 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCCGCTGATAGTGGCCATCTTCTGCTCTGCTG 60
QY 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40
Db 61 GACCTGATGACCGCGCGCAACGCTGGCTGGCTGACGCTACCCACAGGCGCCCTGCGCACTG 120
QY 41 ProGlyLeuGlyAsnLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 121 CCGGGCTGGCAACCTGCTGATGGACTTGGACITTCAGAGACACACCATCTGCTTCGACCA 180
QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValVal 80
Db 181 TTGGCGGCGCGCTTCGGGGACGTTTCAGCTGCGAGCTGGCCTGGACGCGGTGCTCGTG 240
QY 81 LeuAsnGlyLeuAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTCAATGGGCTGGCGCGCGCTGGCGAGGCGCTGGTGAACCCACGCGAGGACACCCCGCAC 300
QY 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln 117
Db 301 CCGCGCGCTGTGCCCCATCACCCAGATCTCTGGGTTTCGGGCGCGCTTCCCAAGGGGTGTT 360
QY 117 ----- 117
Db 361 CTGGCGCGCTATGGCGCGCGTGGCGGAGAGAGAGCGGCTTCTCGTGTCCACCTTGGCG 420
QY 117 ----- 117
Db 421 AACTTTGGGCTGGGCAAGAGTCTGCTGGAGCAGTGGGTGACCGAGGAGCGCGCTTGCCTT 480
QY 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
Db 481 TGTGGCGCGCTTCCCAACACATCCCGAGCGCCCTTTCGCGCCCAACGGTCTCTTTGACAAA 540
QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149
Db 541 GCGGTGAGCAACGTGATCGCCTCCCTCACCTGCGGGCGCGCTTCGAGTACGACGACCTT 600
150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluLysSerGlyPheLeu 169
601 CGCTTCCTCAGCTGCTGGAGCTAGCTCAGAGGAGCTGAAGAGGAGTGGCGCTTTCTG 660
170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
661 CGGAGGTGCTGAATGCTGCTCCCGTCTCTCGCATATCCCGAGCGCTGGTGGCAAGTCT 720
190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
721 CTAGCTTCCAAAAGGCTTCTCTGACCCAGCTGGATGAGTCTGAATCTAACTGACGACAGGATG 780
210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
781 ACCTGGGACCCAGCGCCAGCCCGGACCTGAGCTGAGGGCTTCTCGGAGAGATGGAG 840
230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
841 AAGGCCAAGGGAACTCTGAGAGCAGCTTCAATGATGAGAACCTGGCATAGTGGTGGCT 900
250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 269
901 GACCTGTTCTCTGCGGATGCTGACCACTCGACCGCTCACATGCCCTACACACCTGCGTGTCT 960
270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGlnIleAspAspValIleGly 289
961 ATGATCTTACATCCGATGTCAGCGCGCTGTCACACAGGAGATCGACGATGATAGGG 1020
290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
1021 CAGGTGCGCGCACGAGATGGGTGACGAGCTCACATGCCCTACACACCTGCGTGTATT 1080
310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
1081 CATGAGGTGACGCGCTTTGGGGACATGCTCCCTCGGTATGACCATATGACATCCCGT 1140
330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrLeuIleThrAsnLeuSer 349
1141 GACATCGAAGTACAGGGCTTCCGATCCCTAAGGGAAGCAGCACTCATCACCACTGTCA 1200
350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
1201 TCGTGTCTGAAGATGAGCGCTCTGGAGAGAGCGCTTCCGCTTCCACCCGACACTTC 1260
370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
1261 CTGGATGCCAGGCGCACTTTGTGAAGCGGAGGCTTCTCTGCTTCTCAGCAGGCGCGC 1320
390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuLeuPhePheThrSerLeu 409
1321 CGTGCATGCTTCGGGAGCGCTGCGCGGAGCTTCTGCTGAGTCTTCTCTTCTCACCTCCCTG 1380
410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
1381 CTGACGACTTCAGCTTCTCGGTGCGCCACTGGACAGCCCGCGCCAGCAGCATGCTGTC 1440
430 PheAlaPheLeuValThrProSerProTyrGlu-LeuCysAlaValProArg 446
1441 TTTGCTTCTGCTGGTGACCCCATCCCTTATGAGTATTGTGCTGTCGCCCGC 1492
RESULT 14
AX687030
LOCUS AX687030 1568 bp DNA linear PAT 31-MAR-2003
DEFINITION Sequence 4 from Patent EP1281755.
ACCESSION AX687030
VERSION AX687030.1 GI:29409534
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

AUTHORS Milos, P.M. and Webb, S.M.
TITLE Variants of the human cyd2d6 gene
JOURNAL Patent: EP 1281755-A 4 05-FEB-2003;
Pfizer Products Inc. (US)
FEATURES Location/Qualifiers

source
1..1568

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1,75e-183 Length: 1568
Score: 2267.50 Matches: 442
Percent Similarity: 89.36% Conservative: 3
Best Local Similarity: 88.76% Mismatches: 1
Query Match: 97.32% Indels: 52
DB: 6 Gaps: 1

US-09-820-788A-2 (1-446) x AX687030 (1-1568)

```
Qy 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
Db 1 ATGGGGCTAGAGCACTGGTCCCTGGCGCTGATAGTGGCCATCTTCTGCTCTGGTG 60
Qy 21 AspLeuMetHisArgArgGlnArgTTPAlaAlaArgTyrSerProGlyProLeu 40
Db 61 GACCTGATGACCGCGCCCAACCTGGGCTGCACGTACCACACAGGCCCTGCCACTG 120
Qy 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
Db 121 CCGGGCTGGGCAACCTGCTGCATGTGGACTTCAGAACACACACATCTGCTTCGACCAG 180
Qy 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaIleValValVal 80
Db 181 TTGCGGCGCGCTTCGGGACGTTTTCAGCTCAGCTGCGCTGCGCTGCGGCGGTTGCTG 240
Qy 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
Db 241 CTCAATGGGCTGGCGCGCGCTGGCGAGGCGCTGTCACCCACGCGAGGACACCGCGAC 300
Qy 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln 117
Db 301 CGCGCGCGCTGTGCCCATCACCCAGATCTCTGGGTTTCGGGCGCGCTTCCCAAGGGGTG 360
Qy 117 ----- 117
Db 361 CTGGCGCGCTATGGGCGCGCGTGGCGGACGACAGAGCGCTTCTCCGTGTCCACTTGGCG 420
Qy 117 ----- 117
Db 421 AACTTGGGCTGGGCAAGAGTCTGTCGACAGTGGGTGACCGAGGAGGCGCGCTGCCTT 480
Qy 118 ----- GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
Db 481 TGTGCGCGCTTTCGCCAACCACTCCGGACGCGCTTTCGCCCAACGCTCTCTTGGACAAA 540
Qy 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149
Db 541 GCGGTGAGCAACGTGATGCTCCCTCACCTGCGGCGCGCTTCGAGTACGACGACGCCCT 600
Qy 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGlyPheLeu 169
Db 601 CGCTTCTCAGGCTGTGGACCTAGCTCAGGAGGACTGAAGGAGGAGTGGGCTTCTG 660
Qy 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
Db 661 CGGAGGTGCTGAATGTGTCCCGTCTCTGCTATATCCAGCGCTGGCTGGCAAGGTC 720
Qy 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
Db 721 CTACGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGCTGTAATCTAGCAGCAGGATG 780
```

```
Qy 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
Db 781 ACCTGGGACCCAGCCAGCCCGGAGACCTGATGAGGCTTCTTGGCAGAGATGGAG 840
Qy 230 LysAlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgIleValValAla 249
Db 841 AAGGCCAAGGGGACCTTGAGAGCAGCTTCATGATGAGAACTGGGCATAGTGGTGGCT 900
Qy 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 269
Db 901 GACCTGTTCTCTGCGGGATGGTACCACCTCGACACCGCTGGCGCTGGGCGCTCTCTGCTC 960
Qy 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnGluIleAspAspValIleGly 289
Db 961 ATGATCTTACATCCGGATGTGAGCGCGCTGTCCAACAGAGATCGACACGCTGATAGGG 1020
Qy 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
Db 1021 CAGGTGCGCGGACGAGAGATGGTGACGAGGTTCATGCTCCCTACACCTGCGCGTATT 1080
Qy 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
Db 1081 CATGAGGTGCAGCGCTTTTGGGACATCGTCCCGCTGGGTATGACCCATATGACATCCCGT 1140
Qy 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrThrIleLeuThrAsnLeuSer 349
Db 1141 GACATCGAAGTACAGGGCTTCGCGCATCCCTAAGGGAACGACACTCATCAACCACTGTCA 1200
Qy 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
Db 1201 TCGGTGCTGAAGAGTAGGCGCTCTGGGAGAGCGCTTCCGCTTCCACCCGACACTTC 1260
Qy 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
Db 1261 CTGGATGCCAGGCGCACCTTCTGAAGCGGAGGCTTCTTCCCTTCTCAGCAGCGCGC 1320
Qy 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 409
Db 1321 CGTGATGCTCGGGGAGCGCTTGGCGCGCATGGAGCTTCTCTCTTCTTCTTCTTCTTCT 1380
Qy 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
Db 1381 CTGACGACTTCACTTCTCGGTGCCACTGACAGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Qy 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
Db 1441 TTTGCTTCTCTGAGCGCATCCCGCTATGATGATTGCTGTGCGCGCGC 1492
```

RESULT 15

AR084365
LOCUS AR084365 1566 bp DNA linear PAT 01-SEP-2000
DEFINITION Sequence 13 from patent US 5981174.
ACCESSION AR084365
VERSION AR084365.1 GI:10011136
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1566)
AUTHORS Wolf,C.Roland., Miles,J.Stephen., Spurr,N.Kay. and Gough,A.Charles.
TITLE Genetic assay
JOURNAL Patent: US 5981174-A 13 09-NOV-1999;
FEATURES Location/Qualifiers
1..1566
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2,46e-182 Length: 1566
Score: 2254.00 Matches: 442
Percent Similarity: 89.34% Conservative: 2
Best Local Similarity: 88.93% Mismatches: 2

Query Match: 96.74% Indels: 52
DB: 6 Gaps: 1
US-09-820-788A-2 (1-446) x AR084365 (1-1566)

QY 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
DB 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCCGTGATAGTGGCCATCTTCTCTGCTCTGTTG 60
QY 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40
DB 61 GACCTGATACACGGCGGCAACGCTGGGTGACGCTACCCACCGAGGCCCTTGGCACTG 120
QY 41 ProGlyLeuGlyAsnLeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGln 60
DB 121 CCGGGCTGGGAACCTGCTGATGGACTTCCAGACACACCATACTGCTTCGACCAG 180
QY 61 LeuArgArgArgPheGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValVal 80
DB 181 TTGGCGCGCGCTTCGGGGACGTTTTCAGCCTGCAGCTGGCCTGGACGCGGTGGTCTG 240
QY 81 LeuAsnGlyLeuAlaAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAsp 100
DB 241 CTCAATGGGCTGGCGCGCGCTGGCGAGCGCTGGTGACCCACGGCGAGGACACCGCCGAC 300
QY 101 ArgProProValProIleThrGlnIleLeuGlyPheGlyProArgSerGln----- 117
DB 301 CGCCCGCCCTGTGCCCATCACCCAGATCTCTGGGTTTCGGGCCGCGTTCCCAAGGGGTG 360
QY 117 ----- 117
DB 361 CTGGCGCGCTATGGGCGCGCTGGCGGAGCAGAGCGGCTTCTCGGTGCCACTTGGCG 420
QY 117 ----- 117
DB 421 AACTTGGGCTGGCAAGAGTGGTGGAGCAGTGGTGACCGAGGAGCGCGCTTGCCTT 480
QY 118 -----GlyArgProPheArgProAsnGlyLeuLeuAspLys 129
DB 481 TGTGGCGCTTCGCCNACCACCTCCGGAGCGCCCTTTCGCCCCACAGGTCTCTTGGACAAA 540
QY 130 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPro 149
DB 541 GCCGTGAGCAACGTCATGCCCTCCCTCACCTCGCGGCGCGCTTCGAGTACGACGACCT 600
QY 150 ArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLeu 169
DB 601 CGCTTCTCAGGCTGCTGGACCTAGCTCAGGAGGAGCTGAAGGAGGAGTCTGGGCTTCTG 660
QY 170 ArgGluValLeuAsnAlaValProValLeuLeuHisIleProAlaLeuAlaGlyLysVal 189
DB 661 CGCGAGGTGCTGAATGCTGTCCCGCTCTCTCGCATATCCCAAGCGCTGGTGGCAAGTTC 720
QY 190 LeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuThrGluHisArgMet 209
DB 721 CTACGCTTCCAAAAGGCTTTCCTGACCCAGCTGGATGAGCTGCTAACTGAGCACAGGATG 780
QY 210 ThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeuAlaGluMetGlu 229
DB 781 ACCTGGGACCCAGCCCGCCCGAGACCTGACTGAGGGCTTCTCTGCGAGAGATGGAG 840
QY 230 LysAlaLysGlyAsnProGluSerPheAsnAspGluAsnLeuArgIleValValAla 249
DB 841 AAGGCCAAGGGGACCTTGAGAGCAGCTTCATGATGAGAACCTGCGGCATGATGGTGGCT 900
QY 250 AspLeuPheSerAlaGlyMetValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeu 269
DB 901 GACCTGTTCTCTGCCGGATGGTGACCACTCGACCAAGCTGGCGCTGGGGCTCTCTGCTC 960
QY 270 MetIleLeuHisProAspValGlnArgArgValGlnGlnIleAspAspValIleGly 289
DB 961 ATGATCTCTACATCCGGATGTGCAGC--CGTGTCCAACAGGAGATCGACGACGTGATGGG 1018

QY 290 GlnValArgArgProGluMetGlyAspGlnAlaHisMetProTyrThrThrAlaValIle 309
DB 1019 CAGGTGGCGGACACAGAGATGGGTGACAGGCTACATGCTCCCTACACACCTGCCGTGATT 1078
QY 310 HisGluValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArg 329
DB 1079 CATGAGGTGCAGCGCTTTGGGGACATCGTCCCTGGGTATGACCATATGACATCCCGT 1138
QY 330 AspIleGluValGlnGlyPheArgIleProLysGlyThrLeuIleThrAsnLeuSer 349
DB 1139 GACATCGAAGTACAGGGCTTCGGCATCCCTAAGGGAGACGACTCATCACAACCTGTCA 1198
QY 350 SerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHisProGluHisPhe 369
DB 1199 TCGGTGCTGAGAGATGAGCGCGTCTGGGAGAGCGCTTCCGCTTCCACCCGGAACACTTC 1258
QY 370 LeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPheSerAlaGlyArg 389
DB 1259 CTGGATGCCAGGGCCACTTTGTGAAGCCGGAGGCTTCTCGCTTCTCAGCAGGCCGC 1318
QY 390 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrSerLeu 409
DB 1319 CGTGATCCCTCGGGGAGCCCTGGCCCGCATGGAGCTTCTTCTTCTTCACTCCCTG 1378
QY 410 LeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisGlyVal 429
DB 1379 CTGCAGCACTTCAGCTTCTCGGTGCCCACTGGACAGCCCGGCCAGCCACCATGGTGTTC 1438
QY 430 PheAlaPheLeuValThrProSerProTyrGluLeuCysAlaValProArg 446
DB 1439 TTTGCTTCTCTGGTGAGGCCATCCCTCATGAGCTTTGTGCTGTGCCCGC 1489

Search completed: February 29, 2004, 01:10:51
Job time : 3629 secs


```
FT intron /number= 4
FT /4674..4865
FT /tag= i
FT /number= 4
FT /4866..5007
FT /tag= j
FT /number= 5
FT /replace(4908,T)
FT /tag= s
FT /standard_name= "Single nucleotide polymorphism"
FT /5008..5201
FT /tag= k
FT /number= 5
FT /5202..5389
FT /tag= l
FT /number= 6
FT /5390..5843
FT /tag= m
FT /number= 6
FT /replace(5627,A)
FT /tag= t
FT /standard_name= "Single nucleotide polymorphism"
FT /5844..5985
FT /tag= n
FT /number= 7
FT /5986..9556
FT /tag= o
FT /number= 7
FT /replace(6733,C)
FT /tag= u
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(7788,C)
FT /tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(7867,A)
FT /tag= w
FT /standard_name= "Single nucleotide polymorphism"
FT /replace(7948,T)
FT /tag= x
FT /standard_name= "Single nucleotide polymorphism"
FT /9557..9732
FT /tag= p
FT /number= 8
```

WO200279233-A1.

10-OCT-2002.

01-APR-2002; 2002WO-US009738.

30-MAR-2001; 2001US-00820788.

(PEKE) PE CORP NY.
(DFRA/) DI FRANCESCO V.
(BEAS/) BEASLEY E M.

Shao W, Yan C;

WPI; 2003-040649/03.
P-PSDB; AAO26404.

New human drug-metabolizing proteins and nucleic acids related to the
Cytochrome P450 1D drug-metabolizing enzyme subfamily, useful for
treating a condition mediated by a human enzyme protein e.g., cancer.

Claim 4; Fig 3A-C; 72pp; English.

The invention relates to a novel isolated polypeptide comprising a 446-
amino acid sequence or its allelic variant, orthologue or fragment. The
allelic variant or orthologue is encoded by a nucleic acid that
hybridizes under stringent conditions to the opposite strand of the
nucleic acid comprising a sequence having 1537 or 10278 bp. The fragment
comprises at least 10 contiguous amino acids of the 446-amino acid

CC sequence. The polypeptide is useful for preparing a pharmaceutical
CC composition for treating a disease or condition mediated by a human
CC enzyme protein, e.g. cancer or Parkinson's disease. This polynucleotide
CC sequence represents the genomic DNA encoding the human drug-metabolising
CC protein of the invention

SQ Sequence 10278 BP; 2071 A; 2851 C; 3130 G; 2127 T; 0 U; 99 Other;

Query Match 99.0%; Score 10179; DB 7; Length 10278;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 10278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCCTTACAAAGTGTCTGGATTACCTGGTGAGCCACCGGTCGGGCTCTTTATGTCTT 60

Db 1 AGCCTTACAAAGTGTCTGGATTACCTGGTGAGCCACCGGTCGGGCTCTTTATGTCTT 60

QY 61 ACTGTACTGTCTGCTTGAAGAAGTACTTATTTTGAATGGTTTCATCATTTAGTCTAA 120

Db 61 ACTGTACTGTCTGCTTGAAGAAGTACTTATTTTGAATGGTTTCATCATTTAGTCTAA 120

QY 121 TTAATAAAGAGTAGTTTACACACCAATTTACAGTATTAATACTCTGTTTCTGTG 180

Db 121 TTAATAAAGAGTAGTTTACACACCAATTTACAGTATTAATACTCTGTTTCTGTG 180

QY 181 TGCTTACTATTACAGTGAGTTTGTACCTTTAGATGATTTCTTCTGCTCATTAATATC 240

Db 181 TGCTTACTATTACAGTGAGTTTGTACCTTTAGATGATTTCTTCTGCTCATTAATATC 240

QY 241 CTTTTTTTTTTCAGATTGAAAAAATCCCTTTAGCATTTCTTGTGGATATAGTCTGGTG 300

Db 241 CTTTTTTTTTTCAGATTGAAAAAATCCCTTTAGCATTTCTTGTGGATATAGTCTGGTG 300

QY 301 TTGATGAATCTCGCAGCTTTTGTGTGTCGGGAAGTCTTTATTTCTTCTCAGCATT 360

Db 301 TTGATGAATCTCGCAGCTTTTGTGTGTCGGGAAGTCTTTATTTCTTCTCAGCATT 360

QY 361 AAGGATATTTTGGCAGATACGTTATCTAGCTTAAAGTCTTTTCTTCTCAGCATT 420

Db 361 AAGGATATTTTGGCAGATACGTTATCTAGCTTAAAGTCTTTTCTTCTCAGCATT 420

QY 421 AAATATGTATGCCACTCCCCCTGGCTGTAAAGTCTTCCACTGGAAGTGGTGCCTCC 480

Db 421 AAATATGTATGCCACTCCCCCTGGCTGTAAAGTCTTCCACTGGAAGTGGTGCCTCC 480

QY 481 ATGTCATGTATGGAGCTCTACTGATGTTATTTGTTTCTTCTTCTGCTGCTTTAGG 540

Db 481 ATGTCATGTATGGAGCTCTACTGATGTTATTTGTTTCTTCTTCTGCTGCTTTAGG 540

QY 541 ATCTTTCTTTATCTTGGAGCTTCTGAGTCTTATTAATATCAGATGCTTGGCTCTTC 600

Db 541 ATCTTTCTTTATCTTGGAGCTTCTGAGTCTTATTAATATCAGATGCTTGGCTCTTC 600

QY 601 TTTGGGTTAAATCTCTGTTGGTGTCTATAAACTTTCTGTACAAAAAATCAGCCAGCATG 660

Db 601 TTTGGGTTAAATCTCTGTTGGTGTCTATAAACTTTCTGTACAAAAAATCAGCCAGCATG 660

QY 661 GTGTGGGCACCTGTATATCCAGCTACTTGGAGGCTGAGGAGAGATGCTTTGAACC 720

Db 661 GTGTGGGCACCTGTATATCCAGCTACTTGGAGGCTGAGGAGAGATGCTTTGAACC 720

QY 721 CTGGAGGTGGAGGTTGACGAGTCCGAGATCGCATCTTGGCTCCACCTGGGGGACAG 780

Db 721 CTGGAGGTGGAGGTTGACGAGTCCGAGATCGCATCTTGGCTCCACCTGGGGGACAG 780

QY 781 AGCAAACTCCGCTCTCAAAAAAATAATTTTGGGCTCGGTGGTGTAGTCCAGCT 840

Db 781 AGCAAACTCCGCTCTCAAAAAAATAATTTTGGGCTCGGTGGTGTAGTCCAGCT 840

QY 841 ACTTGGGAGGAGGAGGTCACCTTGAATTTGAGATTTGAGTGGAGTATCTTCCACT 900

Db 841 ACTTGGGAGGAGGAGGTCACCTTGAATTTGAGATTTGAGTGGAGTATCTTCCACT 900

QY 901 GCATCTCCGGCCCGGACACAGAGTGAGACCTGTCTAAAGAAAAAATAAAAAAAG 960

Db	901	GCACCTCCGGCCCGGCAACAGAGTGAGACCTGTCTAAAGAAAAATAAAAAAAG	960	1981	GCTCGGTGTCGAGAGTGCTCTGCTGGTCTCTGTGCTGGTGGGGTGCGGCTGCCAG	2040
Qy	961	CAACATATCTTAATAAAGGATCCTCCATAATGTTTCCACCAGATTTCTAATCAGAAACA	1020	2041	GTGTGTCCAGAGAGCCAGTTGGTAGTGAGGACGACATGGGGCTAGAGACATGGTGCC	2100
Db	961	CAACATATCTTAATAAAGGATCCTCCATAATGTTTCCACCAGATTTCTAATCAGAAACA	1020	2041	GTGTGTCCAGAGAGCCAGTTGGTAGTGAGGACGACATGGGGCTAGAGACATGGTGCC	2100
Qy	1021	TGAGGCCAGGAAGCAGTGAGGAATGACGACCTCAGGCAGGCGCTGGAGATGCTGTAC	1080	2101	CCTGGCCATGATAGTGCCATCTTCTGCTCTGTTGGACCTGATGACACGGGACCAACG	2160
Db	1021	TGAGGCCAGGAAGCAGTGAGGAATGACGACCTCAGGCAGGCGCTGGAGATGCTGTAC	1080	2101	CCTGGCCATGATAGTGCCATCTTCTGCTCTGTTGGACCTGATGACACGGGACCAACG	2160
Qy	1081	AGGCTGGGCAAGGCTTTCAGGCTACCACTGGGAGCTCTGGGAACGCGCTGTTGCAA	1140	2161	CTGGGCTGCACCTACCGCCAGGTCCTCTGCACTGCCCCGGCTGGGCAACCTTCTGTC	2220
Db	1081	AGGCTGGGCAAGGCTTTCAGGCTACCACTGGGAGCTCTGGGAACGCGCTGTTGCAA	1140	2161	CTGGGCTGCACCTACCGCCAGGTCCTCTGCACTGCCCCGGCTGGGCAACCTTCTGTC	2220
Qy	1141	ACAGGAAGTCATGGCCCGGCAGAGCCAGAGATGTGGGCTGAGCTGGGATCCATGTGACA	1200	2221	ATGTGGACTTTCAGAAACACACCATACTGCTTCGACAGGTGAGGGAGGAGGCTCTGGAG	2280
Db	1141	ACAGGAAGTCATGGCCCGGCAGAGCCAGAGATGTGGGCTGAGCTGGGATCCATGTGACA	1200	2221	ATGTGGACTTTCAGAAACACACCATACTGCTTCGACAGGTGAGGGAGGAGGCTCTGGAG	2280
Qy	1201	GCTTTGAGGCTCACCGGGAGCAGCCTCTGGACAGGAGAGTCCATCCAGGAACCTCGG	1260	2281	GGCGCAGAGGTCCTGAGGATGCCCAACACAGGAAACATGGGTGGTGGGTTAAACCA	2340
Db	1201	GCTTTGAGGCTCACCGGGAGCAGCCTCTGGACAGGAGAGTCCATCCAGGAACCTCGG	1260	2281	GGCGCAGAGGTCCTGAGGATGCCCAACACAGGAAACATGGGTGGTGGGTTAAACCA	2340
Qy	1261	GCATGGCTGGGAAGTGGGGTACTTGGTCCGGCTCTGTATGTGTGTGCTGCTGGTGTG	1320	2341	GGCTGGATCAGAAAGCCAGGCTGAGAAAGGGAAGCAGGTTTGGGGACGTTCTTGGGAA	2400
Db	1261	GCATGGCTGGGAAGTGGGGTACTTGGTCCGGCTCTGTATGTGTGTGCTGCTGGTGTG	1320	2341	GGCTGGATCAGAAAGCCAGGCTGAGAAAGGGAAGCAGGTTTGGGGACGTTCTTGGGAA	2400
Qy	1321	TGAGAGAGATGTGCTCCCTGAGTGCAGTGCTGTATGTGTGTAATTTGCTCTT	1380	2401	GACATTTATACATGTCATGAGGACTGGATTTTCCAAAGGCCAAAGGAGTAGGCAAG	2460
Db	1321	TGAGAGAGATGTGCTCCCTGAGTGCAGTGCTGTATGTGTGTAATTTGCTCTT	1380	2401	GACATTTATACATGTCATGAGGACTGGATTTTCCAAAGGCCAAAGGAGTAGGCAAG	2460
Qy	1381	TGTGGGCTGATTTCTGCAATGTGTAATCGTGCTCCCTGCAAGTGTGAACAGTGGACAAG	1440	2461	GGCTGGAGGTGGAGTGGACTTGGCAGTGGGCGATGCAAGCCCATTTGGGCAACATATGT	2520
Db	1381	TGTGGGCTGATTTCTGCAATGTGTAATCGTGCTCCCTGCAAGTGTGAACAGTGGACAAG	1440	2461	GGCTGGAGGTGGAGTGGACTTGGCAGTGGGCGATGCAAGCCCATTTGGGCAACATATGT	2520
Qy	1441	TGTCTGGAGTGACAAAGATCTGTGCACATCAGGTGTGTGTCATAGGCTCTGTGCATG	1500	2521	ATGGAGTACAAAGTCCCTTCTGCTGACACACAGAAAGAAAGGCTTGGGAATGGAAGATGA	2580
Db	1441	TGTCTGGAGTGACAAAGATCTGTGCACATCAGGTGTGTGTCATAGGCTCTGTGCATG	1500	2521	ATGGAGTACAAAGTCCCTTCTGCTGACACACAGAAAGAAAGGCTTGGGAATGGAAGATGA	2580
Qy	1501	TCAAGAGTCAAGGTGAAGTGAAGGACAGGCGCCATGATGCCACATCATCAGGAGCT	1560	2581	GTTAGTCTCAGTGCGGCTTAAATCAACGAAATCGAGGATGAAGGGGTCAGTGACCCCG	2640
Db	1501	TCAAGAGTCAAGGTGAAGTGAAGGACAGGCGCCATGATGCCACATCATCAGGAGCT	1560	2581	GTTAGTCTCAGTGCGGCTTAAATCAACGAAATCGAGGATGAAGGGGTCAGTGACCCCG	2640
Qy	1561	CTAAGGCCCCAGGTAAAGTGCAGTACAGATAAGGGTGTGAAGGTCACCTCTGGAGTGGG	1620	2641	TTCAAACCTTTTGCATGTGGGTCTCTGGGCTCTCACTGCTCACCGGCATGACCATCATC	2700
Db	1561	CTAAGGCCCCAGGTAAAGTGCAGTACAGATAAGGGTGTGAAGGTCACCTCTGGAGTGGG	1620	2641	TTCAAACCTTTTGCATGTGGGTCTCTGGGCTCTCACTGCTCACCGGCATGACCATCATC	2700
Qy	1621	CAGGTGGGGTAAAGGCAAGGTCATGTTCTGGAGAGGGTGTGACTACATTTAG	1680	2701	TGGGAATGGGATGCTAACTGGGGCTCTCGGCAATTTTGGTGAATTTCTGCAAGGTCATAC	2760
Db	1621	CAGGTGGGGTAAAGGCAAGGTCATGTTCTGGAGAGGGTGTGACTACATTTAG	1680	2701	TGGGAATGGGATGCTAACTGGGGCTCTCGGCAATTTTGGTGAATTTCTGCAAGGTCATAC	2760
Qy	1681	GGTGTATGACCTAGCTGGGAGTGGATGSCCGGCTCCATGAGACCTGGTTATCCAG	1740	2761	CTGGGTGACGATCCAACTGAGTTCCTCATCAAGAAAGGTGTGACCCCGCTGCTGCC	2820
Db	1681	GGTGTATGACCTAGCTGGGAGTGGATGSCCGGCTCCATGAGACCTGGTTATCCAG	1740	2761	CTGGGTGACGATCCAACTGAGTTCCTCATCAAGAAAGGTGTGACCCCGCTGCTGCC	2820
Qy	1741	AAGCCTGTGTGGGCTTGGGAGCTTGGAGTGGGAGAGGGGTCATTTCCAGCAGGC	1800	2821	CCACGATCAGGAGGTGGGTCTCTCTTCCACCTGCTCACTCTGGTAGCCCGCGGGGT	2880
Db	1741	AAGCCTGTGTGGGCTTGGGAGCTTGGAGTGGGAGAGGGGTCATTTCCAGCAGGC	1800	2821	CCACGATCAGGAGGTGGGTCTCTCTTCCACCTGCTCACTCTGGTAGCCCGCGGGGT	2880
Qy	1801	CTTTTCAACACCTTACCTGGGTAAAGGCTCTGGAGAGGAGCAGCGGCAAGGACCTCTG	1860	2881	CGTCCAGGTTCAATAGGACTAGGACCTGTAGTCTGGGGTGTGCTGGCTTGCACAGAG	2940
Db	1801	CTTTTCAACACCTTACCTGGGTAAAGGCTCTGGAGAGGAGCAGCGGCAAGGACCTCTG	1860	2881	CGTCCAGGTTCAATAGGACTAGGACCTGTAGTCTGGGGTGTGCTGGCTTGCACAGAG	2940
Qy	1861	GAGCAGCCATACCTGCGGCTGACTGCGCATGCGCAGCAGTCAACACAGCAGG	1920	2941	GGCTGACCTCTCTCTGAGTTGGCGCGCTTCCGGGACGTTTTCAGGCTTGCAGCTG	3000
Db	1861	GAGCAGCCATACCTGCGGCTGACTGCGCATGCGCAGCAGTCAACACAGCAGG	1920	2941	GGCTGACCTCTCTCTGAGTTGGCGCGCTTCCGGGACGTTTTCAGGCTTGCAGCTG	3000
Qy	1921	TTCACTCAGCAGAGGCGAAGCCATCATCAGTCCCTTTTAAAGGGAAGGGTCAAGC	1980	3001	GGCTGGAACCGCGGTGCTGCTCAATAGGCTGGCGGCTGCGCAGGCGATGGTGACC	3060
Db	1921	TTCACTCAGCAGAGGCGAAGCCATCATCAGTCCCTTTTAAAGGGAAGGGTCAAGC	1980	3001	GGCTGGAACCGCGGTGCTGCTCAATAGGCTGGCGGCTGCGCAGGCGATGGTGACC	3060
Qy	1981	GCTCGGTGTCGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2040	3061	CGGGGAGGACACGGCCGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG	3120
Db	1981	GCTCGGTGTCGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2040	3061	CGGGGAGGACACGGCCGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG	3120

Db 5281 CACCATGCGCGTGAATCACGAGGTGCGAGCGCTTTGGGACATCATCCCCCTGAGTGTGAC 5340
Qy
Db 5341 CCATATGACATCCCGTGACATCGAAGTACAGGGCTTCGCGATCCCTAAGGTAGGCTTGGC 5400
Db 5341 CCATATGACATCCCGTGACATCGAAGTACAGGGCTTCGCGATCCCTAAGGTAGGCTTGGC 5400
Qy 5401 GCCCTCTCACCCAGCTCAGCACAGCAGCACTGTGTATAGCCCGAGCATGGCTACTGCCA 5460
Db 5401 GCCCTCTCACCCAGCTCAGCACAGCAGCACTGTGTATAGCCCGAGCATGGCTACTGCCA 5460
Qy 5461 GGTGGGCCCACTCTAGGAACCTCGGCCACTAGTCTCAATAGCCACACACTGACTGTCC 5520
Db 5461 GGTGGGCCCACTCTAGGAACCTCGGCCACTAGTCTCAATAGCCACACACTGACTGTCC 5520
Qy 5521 CCATCTGGGTGGGGGTCCAGAGTATAGGCAGGGCTGGCCCTGTCCATCCAGAGCCCCCGT 5580
Db 5521 CCATCTGGGTGGGGGTCCAGAGTATAGGCAGGGCTGGCCCTGTCCATCCAGAGCCCCCGT 5580
Qy 5581 CTAGTGGGGAGACAAACAGGACCTGCCAGAAATGTTGGAGGCCAGCGCTCGAGGGAG 5640
Db 5581 CTAGTGGGGAGACAAACAGGACCTGCCAGAAATGTTGGAGGCCAGCGCTCGAGGGAG 5640
Qy 5641 AGGGGCGAGTGTGGGTGCTCTGAGAGGTGTGACTGCGCCCTGCTGTGGGGTGGAGAGG 5700
Db 5641 AGGGGCGAGTGTGGGTGCTCTGAGAGGTGTGACTGCGCCCTGCTGTGGGGTGGAGAGG 5700
Qy 5701 GTACTGTGGAGCTTCTCGGGCGAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAGGCGAG 5760
Db 5701 GTACTGTGGAGCTTCTCGGGCGAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAGGCGAG 5760
Qy 5761 TGTGTGTCCCGGTGTGTGGTGGCAGGGGTCCAGCATCTCTAGAGTCAGTCCCACT 5820
Db 5761 TGTGTGTCCCGGTGTGTGGTGGCAGGGGTCCAGCATCTCTAGAGTCAGTCCCACT 5820
Qy 5821 CTCACCTGATCTCTGCGCCAGGAGACGACATCTCATCACCACTGTCTATCGGTGCTGA 5880
Db 5821 CTCACCTGATCTCTGCGCCAGGAGACGACATCTCATCACCACTGTCTATCGGTGCTGA 5880
Qy 5881 AGGATGAGGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCCGAACACTTCTCGATGCC 5940
Db 5881 AGGATGAGGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCCGAACACTTCTCGATGCC 5940
Qy 5941 AGGGCCACTTTGTGAAGCGGAGCCCTTCTGCTTTCTCAGAGGTGCTGTGGGAGC 6000
Db 5941 AGGGCCACTTTGTGAAGCGGAGCCCTTCTGCTTTCTCAGAGGTGCTGTGGGAGC 6000
Qy 6001 CCGGCTCCCTGTCCCTTCCGTGGAGTCTTGCAAGGGGTATCACCCAGAGCCAGGCTCAC 6060
Db 6001 CCGGCTCCCTGTCCCTTCCGTGGAGTCTTGCAAGGGGTATCACCCAGAGCCAGGCTCAC 6060
Qy 6061 TGACGCCCTCCCTCCCCACAGCGCGGTGATGCTCGGGAGCCCTCGCCCGCAT 6120
Db 6061 TGACGCCCTCCCTCCCCACAGCGCGGTGATGCTCGGGAGCCCTCGCCCGCAT 6120
Qy 6121 GGAGCTCTTCTCTTCTTCACTCTCTGCTGACGACTTTCAGCTTCTCGGTGGCCCGCG 6180
Db 6121 GGAGCTCTTCTCTTCTTCACTCTCTGCTGACGACTTTCAGCTTCTCGGTGGCCCGCG 6180
Qy 6181 AAGCCCGGCGCCAGGCACTCTCGTGTGTGAGTCTTGTGACCCCATTCGCCCTACGA 6240
Db 6181 AAGCCCGGCGCCAGGCACTCTCGTGTGTGAGTCTTGTGACCCCATTCGCCCTACGA 6240
Qy 6241 GCTTTGTGTGTGCCCGGTAGAAATGGGGTACCTAGTCCCGAGCGCTGCTCCCTAGCCAGA 6300
Db 6241 GCTTTGTGTGTGCCCGGTAGAAATGGGGTACCTAGTCCCGAGCGCTGCTCCCTAGCCAGA 6300
Qy 6301 GGCTCTAAATGTAACAATAAGCAATGTGGTAGTTCCAACTTGGGTCCCTGCTCACGCCCT 6360
Db 6301 GGCTCTAAATGTAACAATAAGCAATGTGGTAGTTCCAACTTGGGTCCCTGCTCACGCCCT 6360
Qy 6361 CGTTGGGATCATCTCTCAGGGCAACCCCAACCCCTGCTCATCTCTGCTTACCCCAACCG 6420

Db 6361 CGTTGGGATCATCTCTCAGGGCAACCCCAACCCCTGCTCATCTCTGCTTACCCCAACCG 6420
Qy
Db 6421 CCTGGCGCATTTGAGACCGGTACGTTGAGGCTGAGCAGATGTCACTTACCTTGGCCAT 6480
Db 6421 CCTGGCGCATTTGAGACCGGTACGTTGAGGCTGAGCAGATGTCACTTACCTTGGCCAT 6480
Qy 6481 AATCCCATGTCCCCCACTGACCCCACTCTGACTGCCAGATTGGTGGACAAGGACTACATT 6540
Db 6481 AATCCCATGTCCCCCACTGACCCCACTCTGACTGCCAGATTGGTGGACAAGGACTACATT 6540
Qy 6541 GTCCTGSCATGTGGGGAAGGGGCCAGAAATGGGCTGACTAGAGGTGTCACTGAGCCCTGGA 6600
Db 6541 GTCCTGSCATGTGGGGAAGGGGCCAGAAATGGGCTGACTAGAGGTGTCACTGAGCCCTGGA 6600
Qy 6601 TGTGGTGGAGAGGCGGAGGACTCAGAGCTGGAGGCCATATTTCAAGGCTTAATCAGCCAC 6660
Db 6601 TGTGGTGGAGAGGCGGAGGACTCAGAGCTGGAGGCCATATTTCAAGGCTTAATCAGCCAC 6660
Qy 6661 CCCACATCAGGGACAGCAGTCTCTGCCAGCACCATCAACAGTCACTCCCTTCATATAT 6720
Db 6661 CCCACATCAGGGACAGCAGTCTCTGCCAGCACCATCAACAGTCACTCCCTTCATATAT 6720
Qy 6721 GACACCCCAAAATGGAAGACAAATCATGTCAAGGACTATATGCCAGGGCTACTCCTCCAG 6780
Db 6721 GACACCCCAAAATGGAAGACAAATCATGTCAAGGACTATATGCCAGGGCTACTCCTCCAG 6780
Qy 6781 GGTCTAGTCCGAGGTGCCAGAACATTTCCCTGGAGGGCCCCAGGAAACCCAGGACCGA 6840
Db 6781 GGTCTAGTCCGAGGTGCCAGAACATTTCCCTGGAGGGCCCCAGGAAACCCAGGACCGA 6840
Qy 6841 GCCACGCCCTCAGCGCTGTCACTTGTGTCCAAATTTGGTGGTTCCTTGTCTCACTGAC 6900
Db 6841 GCCACGCCCTCAGCGCTGTCACTTGTGTCCAAATTTGGTGGTTCCTTGTCTCACTGAC 6900
Qy 6901 TTCAAGAAATGAAGCGGTGACCCCTCAAGGTGAGTGTACAGTTCCTTAAAGATGGTGTGT 6960
Db 6901 TTCAAGAAATGAAGCGGTGACCCCTCAAGGTGAGTGTACAGTTCCTTAAAGATGGTGTGT 6960
Qy 6961 CAGAGTTTGTCTCTCTGATGTGAAGCGTGTTCAGAGTTCCTTCTCTGGTGGGTGG 7020
Db 6961 CAGAGTTTGTCTCTCTGATGTGAAGCGTGTTCAGAGTTCCTTCTCTGGTGGGTGG 7020
Qy 7021 TGGTCTTGTGCTTTCAGGAGTGAAGCTGCAGACCTTCAAGTGAAGTGTACGGCTCTTA 7080
Db 7021 TGGTCTTGTGCTTTCAGGAGTGAAGCTGCAGACCTTCAAGTGAAGTGTACGGCTCTTA 7080
Qy 7081 AGGCTGCAGTACGAGGTGTTCATCTTCTGGTGGGTGTGGTGTCACTGGCCCTCAG 7140
Db 7081 AGGCTGCAGTACGAGGTGTTCATCTTCTGGTGGGTGTGGTGTCACTGGCCCTCAG 7140
Qy 7141 GAGTGAACCTGAGTCCCTTCCAGTGTTCACCTCAATAAGGAGTGTGGACCCCAATGAGG 7200
Db 7141 GAGTGAACCTGAGTCCCTTCCAGTGTTCACCTCAATAAGGAGTGTGGACCCCAATGAGG 7200
Qy 7201 GAGCAGCAGCAGCAAGACTTACTGCACAAACAGCAAAAGAAATGATGCAACAGGTTGCCCG 7260
Db 7201 GAGCAGCAGCAGCAAGACTTACTGCACAAACAGCAAAAGAAATGATGCAACAGGTTGCCCG 7260
Qy 7261 TGCTACTTTCAGGAGCTGTCTTATTCCTTATCTGACCCCAACCCACATCTCTGAT 7320
Db 7261 TGCTACTTTCAGGAGCTGTCTTATTCCTTATCTGACCCCAACCCACATCTCTGAT 7320
Qy 7321 TGGCCCATTTTACAGACAGTGGTTCACCTTACAGAGACTGATTTGGTGCATTTACA 7380
Db 7321 TGGCCCATTTTACAGACAGTGGTTCACCTTACAGAGACTGATTTGGTGCATTTACA 7380
Qy 7381 ATCCCTGAGCTAGACACAGAGTACTGATTTGATATTTTCAAACTTGAAGTACACAG 7440
Db 7381 ATCCCTGAGCTAGACACAGAGTACTGATTTGATATTTTCAAACTTGAAGTACACAG 7440
Qy 7441 AGTGTGAATGGTGTATTTACAAATCCCTTGTAGACATAAAGGTGTGCCAGTCCCCAC 7500
Db 7441 AGTGTGAATGGTGTATTTTAAATCCCTTGTAGACATAAAGGTGTGCCAGTCCCCAC 7500

QY	7501	TAGATTAGCTAGATAGATAGACAGAGACATGATTGGTGGCTTACAAACCTTGAGTT	7560
DB	7501	TAGATTAGCTAGATAGATAGACAGAGACATGATTGGTGGCTTACAAACCTTGAGTT	7560
QY	7561	AGACACAGGGTGTGACTGGTGTGTTTACAAACCTTGAGCTAGACACAGAGTGTGATTG	7620
DB	7561	AGACACAGGGTGTGACTGGTGTGTTTACAAACCTTGAGCTAGACACAGAGTGTGATTG	7620
QY	7621	GTGATTATTAACAATCTTTTGTAGCTAGAAATAAGGTTCCCAAGTCCCAACAGATTAGCTA	7680
DB	7621	GTGATTATTAACAATCTTTTGTAGCTAGAAATAAGGTTCCCAAGTCCCAACAGATTAGCTA	7680
QY	7681	GATAGAGTCTAATTTGGTGCATGCACGAACCCGGAGCTAGACACAGAGTGTGATTGGTG	7740
DB	7681	GATAGAGTCTAATTTGGTGCATGCACGAACCCGGAGCTAGACACAGAGTGTGATTGGTG	7740
QY	7741	CATATACAAATCTCTGGCTAGACATATAAAGTTCTCCAAAGTCCCACTGACTCAGAGCC	7800
DB	7741	CATATACAAATCTCTGGCTAGACATATAAAGTTCTCCAAAGTCCCACTGACTCAGAGCC	7800
QY	7801	CAGCCAGCTTCGCTAGTGGATCTTATGCCCAGGGCCACAGGACAGAGTGCCTGTAGTCC	7860
DB	7801	CAGCCAGCTTCGCTAGTGGATCTTATGCCCAGGGCCACAGGACAGAGTGCCTGTAGTCC	7860
QY	7861	CACACGGGCACTGCTACTCTCAGCCCTTGGGCGAGTGACACAGGTCGCTGGAG	7920
DB	7861	CACACGGGCACTGCTACTCTCAGCCCTTGGGCGAGTGACACAGGTCGCTGGAG	7920
QY	7921	CAGTGGAGGCAACCAATCCGGAGGCTCGGGCTCGCAGGAGCCACCGTAGGAGGCT	7980
DB	7921	CAGTGGAGGCAACCAATCCGGAGGCTCGGGCTCGCAGGAGCCACCGTAGGAGGCT	7980
QY	7981	TGGGCATGCGAGGCTGCAAGTCTTGAGCCCTCCCGCGGGGAGTGACTGAGGCTGCG	8040
DB	7981	TGGGCATGCGAGGCTGCAAGTCTTGAGCCCTCCCGCGGGGAGTGACTGAGGCTGCG	8040
QY	8041	GACAAATCAAGTGTGGTGGCCGGCAGGCGCAGCAGTACTTGGGGACCCGGTGGCCCT	8100
DB	8041	GACAAATCAAGTGTGGTGGCCGGCAGGCGCAGCAGTACTTGGGGACCCGGTGGCCCT	8100
QY	8101	CTGCACTCTCGGCCAGAGTGTAAAGCCCTCACTGCTCGGGCCAGAGGCAACAGCCGG	8160
DB	8101	CTGCACTCTCGGCCAGAGTGTAAAGCCCTCACTGCTCGGGCCAGAGGCAACAGCCGG	8160
QY	8161	CGCTCCGAGTGAGGGCCCGTGTAGCCCTCCGCCACCCAGAACTGGTGTGGCCCGGA	8220
DB	8161	CGCTCCGAGTGAGGGCCCGTGTAGCCCTCCGCCACCCAGAACTGGTGTGGCCCGGA	8220
QY	8221	GCAACCCAGGTTCCGCGACAGCCCTCTCCCTCCATACCTCCCGCAAGCAGAGCGG	8280
DB	8221	GCAACCCAGGTTCCGCGACAGCCCTCTCCCTCCATACCTCCCGCAAGCAGAGCGG	8280
QY	8281	GCTCCAGCTCCACAGTCCAGAGGGGCTCCACAGTGCAGCGCTGGGCTGAACAAGG	8340
DB	8281	GCTCCAGCTCCACAGTCCAGAGGGGCTCCACAGTGCAGCGCTGGGCTGAACAAGG	8340
QY	8341	TCTTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGAGTGTAACTGAGCAGAGA	8400
DB	8341	TCTTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGAGTGTAACTGAGCAGAGA	8400
QY	8401	TGACCTGGGACCCAGCCAGCCCTCCCTCCATACCTCCCGCAAGCAGAGATG	8460
DB	8401	TGACCTGGGACCCAGCCAGCCCTCCCTCCATACCTCCCGCAAGCAGAGATG	8460
QY	8461	GAGAAGGTGAGAGTGGCTGCCACGGTGGGGGCAAGGTTGGGTTGAGCGTCCAGGA	8520
DB	8461	GAGAAGGTGAGAGTGGCTGCCACGGTGGGGGCAAGGTTGGGTTGAGCGTCCAGGA	8520
QY	8521	GGAATGAGGGAGGCTGGGCAAAAGGTTGACAGTGCATCACCCGGGAGCCGATCTG	8580
DB	8521	GGAATGAGGGAGGCTGGGCAAAAGGTTGACAGTGCATCACCCGGGAGCCGATCTG	8580

QY	8581	GSCTGACAGGTGCAGAAATTGGAGGTCAATTTGGGGGTACCCCGTTCTGTCCGAGTATGC	8640
DB	8581	GSCTGACAGGTGCAGAAATTGGAGGTCAATTTGGGGGTACCCCGTTCTGTCCGAGTATGC	8640
QY	8641	TCTCGGCCCTGTCTCAGGCCAAGGGGAAACCTTGAGAGCAGCTTCAATGATGAGAACTCGG	8700
DB	8641	TCTCGGCCCTGTCTCAGGCCAAGGGGAAACCTTGAGAGCAGCTTCAATGATGAGAACTCGG	8700
QY	8701	CATAGTGTGGTGTACCTTGTCTCTCGCGGATGTGACCACTCTGACACAGCTGGGCTG	8760
DB	8701	CATAGTGTGGTGTACCTTGTCTCTCGCGGATGTGACCACTCTGACACAGCTGGGCTG	8760
QY	8761	GSGCCCTCTGTCTCATGATCCCTACATCCCGATGTGAGCGTGAGCCATCTGGGAAACAGT	8820
DB	8761	GSGCCCTCTGTCTCATGATCCCTACATCCCGATGTGAGCGTGAGCCATCTGGGAAACAGT	8820
QY	8821	GCAGGGGCGGAGGGAAGGTTACAGCGGGGGCCCATGAACTTTGCTGGGACACCCGG	8880
DB	8821	GCAGGGGCGGAGGGAAGGTTACAGCGGGGGCCCATGAACTTTGCTGGGACACCCGG	8880
QY	8881	GGCTCCAAAGCAAGGCTTTGACAGGATCTCTGTAAGCTGACCTCTCCCAACATAGAGGC	8940
DB	8881	GGCTCCAAAGCAAGGCTTTGACAGGATCTCTGTAAGCTGACCTCTCCCAACATAGAGGC	8940
QY	8941	AAGAGGAGTGTCAAGGCGCGGACCCCTGGGTGTGACCCATTTGTTGGGACGCGATGCTG	9000
DB	8941	AAGAGGAGTGTCAAGGCGCGGACCCCTGGGTGTGACCCATTTGTTGGGACGCGATGCTG	9000
QY	9001	TCCAGGCCCTGTCCAAACAGAGAGATCGACGCTGTATAGGCGAGGTGCGCGACACAGAT	9060
DB	9001	TCCAGGCCCTGTCCAAACAGAGAGATCGACGCTGTATAGGCGAGGTGCGCGACACAGAT	9060
QY	9061	GGGTGACAGGCTCATGCTCCCTACACCACTGCGGTGATTCATGAGTGCAGCGCTTTGG	9120
DB	9061	GGGTGACAGGCTCATGCTCCCTACACCACTGCGGTGATTCATGAGTGCAGCGCTTTGG	9120
QY	9121	GGAATCGTCCCTGGGTGTGACCACTATGACATCCCGTGCATTCGAACTACAGGCT	9180
DB	9121	GGAATCGTCCCTGGGTGTGACCACTATGACATCCCGTGCATTCGAACTACAGGCT	9180
QY	9181	TCCGATCCCTAAGTAGGCTTGGCGCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	9240
DB	9181	TCCGATCCCTAAGTAGGCTTGGCGCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	9240
QY	9241	NN	9300
DB	9241	NN	9300
QY	9301	NNNNNNCTCCAGGGAAACGACTCATCAACCACTGTCTCGGTGCTGAAGATGA	9360
DB	9301	NNNNNNCTCCAGGGAAACGACTCATCAACCACTGTCTCGGTGCTGAAGATGA	9360
QY	9361	GCGCTGTGGGAGAGCCCTTCCGCTTCCACCCCGAACACTTCTTGGATGCCAGGGCA	9420
DB	9361	GCGCTGTGGGAGAGCCCTTCCGCTTCCACCCCGAACACTTCTTGGATGCCAGGGCA	9420
QY	9421	CTTTGTGAAGCGGAGGCTTCTGCTTCTCAGAGGTGCTGTGGGAGCCCGGCTC	9480
DB	9421	CTTTGTGAAGCGGAGGCTTCTGCTTCTCAGAGGTGCTGTGGGAGCCCGGCTC	9480
QY	9481	CCTGTCCCTTCCGTGGAGTCTTCAGGGGTATCACCCAGGAGCCAGGCTCACTGACGCC	9540
DB	9481	CCTGTCCCTTCCGTGGAGTCTTCAGGGGTATCACCCAGGAGCCAGGCTCACTGACGCC	9540
QY	9541	CCTCCCTTCCACAGGCGCGCTGCTGCTCCTCGGGAGCCCTGGCCCGCATGGAGCTC	9600
DB	9541	CCTCCCTTCCACAGGCGCGCTGCTGCTCCTCGGGAGCCCTGGCCCGCATGGAGCTC	9600
QY	9601	TTCCTTCTTTCACCTCCTGTGAGCACTTCAGCTTCTCGGTGCCCATGGACAGGCC	9660
DB	9601	TTCCTTCTTTCACCTCCTGTGAGCACTTCAGCTTCTCGGTGCCCATGGACAGGCC	9660
QY	9661	CGGCCAGCCACCATGGTGTCTTTGCTTCTGGTGGAGCCCATCCCGCTATGAGCTTTGT	9720

QY 5655 GTGCTCTCAGAGGTGTGACTGGCCCTGCTGTGGGTGGAGAGGTACTGTGGAGCTT 5714
DB 5229 GTGCTCTCAGAGGTGTGACTGGCCCTGCTGTGGGTGGAGAGGTACTGTGGAGCTT 5288
QY 5715 CTGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAGGCAAGTGTGTCCCCCGT 5774
DB 5289 CTGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAGGCAAGTGTGTCCCCCGT 5348
QY 5775 GTGTTTGTGGCAGGGGTCCCAGCATCTTAGAGTCCAGTCCCACTCTCACCCCTGCATCT 5834
DB 5349 GTGTTTGTGGCAGGGGTCCCAGCATCTTAGAGTCCAGTCCCACTCTCACCCCTGCATCT 5408
QY 5835 CTGCCCAGGGAACGACATCATCAACCTGTGTATGGTGTGTGAGGATGAGGCGGTG 5894
DB 5409 CTGCCCAGGGAACGACATCATCAACCTGTGTATGGTGTGTGAGGATGAGGCGGTG 5468
QY 5895 TGGGAGAACCCCTTCGCTTCCACCCCGAACTCTCTGATGCCAGGCGCACCTTTGTG 5954
DB 5469 TGGGAGAACCCCTTCGCTTCCACCCCGAACTCTCTGATGCCAGGCGCACCTTTGTG 5528
QY 5955 AAGCCGAGGCTTCTGCTTTCTCAGCAGTGTGTGGGAGCCCGGCTCCCTGTGC 6014
DB 5529 AAGCCGAGGCTTCTGCTTTCTCAGCAGTGTGTGGGAGCCCGGCTCCCTGTGC 5588
QY 6015 CTTTCGCTGAGTCTTGAGGGGTATCACCCAGGAGCCAGGCTCACTGACGCCCTCC 6074
DB 5589 CTTTCGCTGAGTCTTGAGGGGTATCACCCAGGAGCCAGGCTCACTGACGCCCTCC 5648
QY 6075 TCCCAACAGCCCGCTGCATGCTCGGGAGCCCTGCGCCGCGATGAGGCTTCTCTCT 6134
DB 5649 TCCCAACAGCCCGCTGCATGCTCGGGAGCCCTGCGCCGCGATGAGGCTTCTCTCT 5708
QY 6135 TCTTCACTTCTGCTGACGACTTTCAGCTTCTCGGTGCCACTGACAGCCCGGCCCA 6194
DB 5709 TCTTCACTTCTGCTGACGACTTTCAGCTTCTCGGTGCCACTGACAGCCCGGCCCA 5768
QY 6195 GCACTCTGCTGCTGACGCTTCTGCTGACCCGATCCCGCTAGAGCTTGTGTGTGC 6254
DB 5769 GCACTCTGCTGCTGCTTCTTCTGCTGAGCCATCCCGCTATGAGCTTGTGTGTGC 5828
QY 6255 CCGCTAGAAATGGGTACCTAGTCCCGAGCTGCTCCCTAGCAGAGGCTCTAAAGTACA 6314
DB 5829 CCGCTAGAAATGGGTACCTAGTCCCGAGCTGCTCCCTAGCAGAGGCTCTAAAGTACA 5888
QY 6315 ATAAAGCAATGTGTAGTTCCAACTTGGGTCCCTGCTCAAGCCCTCGTTGGGATCATCC 6374
DB 5889 ATAAAGCAATGTGTAGTTCCAACTTGGGTCCCTGCTCAAGCCCTCGTTGGGATCATCC 5948
QY 6375 TCTCAGGGCAACCCGACCCCTGCTCATCTGCTTACCCAGCCGCTGGCCGATTTG 6434
DB 5949 TCTCAGGGCAACCCGACCCCTGCTCATCTGCTTACCCAGCCGCTGGCCGATTTG 6008
QY 6435 AGAC--GGGTAGTGTGAGCTCAGAGATGTCAGTTACCTTGGCCCAATCCCATGTC 6492
DB 6009 AGACAGGGTACGTTGAGCTCAGAGATGTCAGTTACCTTGGCCCAATCCCATGTC 6068
QY 6493 CCCACTGACCCCACTCTGACTGCCAGATTTGGTGACAGGACTACATTTCTCGGCACTG 6552
DB 6069 CCCACTGACCCCACTCTGACTGCCAGATTTGGTGACAGGACTACATTTCTCGGCACTG 6128
QY 6553 GGGGAAGGGCCAGAAATGGGCTGACTAGAGTGTGATGAGCCCTGGATGTGGTGAGAG 6612
DB 6129 GGGGAAGGGCCAGAAATGGGCTGACTAGAGTGTGATGAGCCCTGGATGTGGTGAGAG 6188
QY 6613 GCGAGACTCAGCTGGAGGCCCATATTTTACGCGCTAATCTCAGCCACCCCATCAGG 6672
DB 6189 GCGAGACTCAGCTGGAGGCCCATATTTTACGCGCTAATCTCAGCCACCCCATCAGG 6248
QY 6673 ACAGCAGTCTCGCCAGCACCATCACAACTGACCTCCCTTATATATGACACCCCAAAA 6732
DB 6249 ACAGCAGTCTCGCCAGCACCATCACAACTGACCTCCCTTATATATGACACCCCAAAA 6308

QY 6733 TGGAGACAAATCAT---GTCAAGGAGCTATATGCAGG---CTACTCCAGGGCTC 6785
DB 6309 CGGAAGACAAATCATGGCGTCAGGAGCTATATGCAGGGCTACTACTCCAGGGCTC 6368
QY 6786 AGTCGCGCAGGTGCGCAAGACATTTCCCTGGGAAGGCCCCAGGAAACCCAGGACCCAGCCAC 6845
DB 6369 AGTCGCGCAGGTGCGCAAGACATTTCCCTGGGAAGGCCCCAGGACCCAGGACCCAGCCAC 6428
QY 6846 CGCCCTCAGCCT 6857
DB 6429 CACCCTCAGCCT 6440

RESULT 3
ACAG1301
ID ACAG1301 standard; DNA; 9432 BP.
XX AC ACAG1301;
XX 16-JUL-2003 (first entry)
XX Human cytochrome p450 gene CYP2D6, wild-type.
XX Human; ds; gene; cytochrome P450; CYP2D6; chromosome 22; SNP;
XX single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
XX psychiatric disorder; drug sensitivity.
XX Homo sapiens.

Key Location/Qualifiers
FT variation replace(226..227,ATT)
FT variation /*tag= a
FT variation /standard_name= "Single nucleotide polymorphism"
FT variation replace(971,G)
FT variation /*tag= b
FT variation /standard_name= "Single nucleotide polymorphism"
FT variation replace(1111,T)
FT variation /*tag= c
FT variation /standard_name= "Single nucleotide polymorphism"
FT variation replace(1726,C)
FT variation /*tag= d
FT variation /standard_name= "Single nucleotide polymorphism"
FT variation replace(1846,A)
FT variation /*tag= f
FT variation /standard_name= "Single nucleotide polymorphism"
FT variation replace(1846,G)
FT variation /*tag= e
FT variation /standard_name= "Single nucleotide polymorphism"
FT variation replace(2064,A)
FT variation /*tag= g
FT variation /standard_name= "Single nucleotide polymorphism"
FT variation replace(3023,A)
FT variation /*tag= h
FT variation /standard_name= "Single nucleotide polymorphism"
FT variation replace(5799,C)
FT variation /*tag= i
FT variation /standard_name= "Single nucleotide polymorphism"
FT variation replace(5816,TA)
FT variation /*tag= j
FT variation /standard_name= "Single nucleotide polymorphism"
XX EPI281755-A2.
XX 05-FEB-2003.
XX 16-JUL-2002; 2002EP-00254972.
XX 31-JUL-2001; 2001US-0309111P.
XX (PFIZ) PFIZER PROD INC.
XX Milos PM, Webb SM;

WPI; 2003-373769/36.

New cytochrome P450 2D6 gene variants and polypeptides, useful for determining if a subject has or is at risk of developing a drug sensitivity condition or disorder that is associated with an aberrant CYP2D6 activity.

Claim 1; Fig 2; 88pp; English.

The invention relates to an isolated nucleic acid comprising a cytochrome P450 2D6 gene variant, e.g. G5799C or C5816A* (referring to the genomic sequence or the same variant nucleotide in the corresponding cDNA sequences). Also included are probes, primers (allele specific oligonucleotides) and arrays used to detect and/or amplify the CYP2D6 gene polymorphic regions, the variant polypeptides, antibodies which are capable of distinguishing between the variant and wild-type polypeptides, determining whether a subject has a genetic deficiency for metabolising a drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and determining whether an individual is susceptible to being a poor metaboliser of drugs. The DNA probe is useful for hybridising to a variant form of the CYP2D6 gene. The primer is useful for amplifying the C5816A* allelic variant. The allele specific nucleotide is useful for the detection of the C5816A* allelic variant. The methods are useful for determining whether a subject has a genetic deficiency for metabolising a drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and determining if an individual is susceptible to being a poor metaboliser of drugs. The nucleic acids are useful as probes or primers for determining whether a subject has a genetic deficiency for metabolising drugs that are substrates of P450 CYP2D6. The methods are useful for determining if a subject has or is at risk of developing a drug sensitivity condition or disorder that is associated with an aberrant CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an aberrant CYP2D6 bioactivity. The methods are also useful in selecting the appropriate drugs or determining the course of treatment to administer to a subject to treat cardiovascular or psychiatric disorders, or for a subject with a drug sensitivity or disorder associated with a specific allelic variant of a polymorphic region of the CYP2D6 gene. The antibodies are useful for monitoring CYP2D6 protein levels in an individual for determining whether a subject has a disease or condition associated with an aberrant CYP2D6 protein level. The gene is located on human chromosome 22. The present sequence is the wild-type CYP2D6 gene

Sequence 9432 BP; 1964 A; 2647 C; 2976 G; 1845 T; 0 U; 0 Other;

51.8%: Score 5325.6: DB 7: Length 9432:

```

Very Match          31.8%, score 3323.6; DB /; length 3432;
1st Local Similarity 93.4%; Pred. No. 0:

```

Percent local similarity	93.4%	Fied: NO: 0;	
Matches 5895; Conservative	0; Mismatches 269;	Indels 148;	Gaps 26;

637 TGTACAAAAAATCAGCCAGGCATGGTGGTGGGCACCTGTAAATCCAGCTACTTGGGAGGC 696

057 1GTACATTTTTCAGCCAGGCGAAGTGCTGGGCACCCTGTAACTCCAGCTACTGGAGGC 898

186 TGAAAAATACAAAAAGCTAGACGTGGTGGCACACACCTGTAAATCCCAGCTACTTAGGAGGC 245

[illegible]

697 TGAGGCAGGAGAAATCGCTTGAACCTGGAGGTGGAGGTGCAGTGAGCCGAGATCGCATC 756

100

246 TGAGGCAGGAGAAATTGCTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCCGAGATTGCATC 305

.....

757 ATTGCA-----CTCCACCTGGGCGAC-AGAGCAAAACTCCGTCTC----- 796

[illegible]

306 ATTGCACAATGGAGGGAGCCACCAGCCTGGGCAACAAGAGGAAATCTCCGTCTCCAAA 365

797 -----AAAAAAATTATTTGGGCTCGGTGGTGCCTGTAGTCCAGCTACTTGGGAG 849

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

366 AAAAAAAAAAAAAAAAAAGAAATTAGGCTGGGTGGTGCTGTAGTCCAGCTACTTGGAG 425

100

850 GCAGGAGGTCCACTTGATGTTGAGATTGCAGTGAGCCATGATCCTGCCACTGCACCTCCGG 909

[illegible]

426 GCAGGGGTCCACTTIGATGTCGAGACTGCAGTGAGCCATGATCCTGCCACTGCACCTCCG 485

910 CCCCCCACAAGACGCTCACGCCCTGTCTAACACAAATATAAATATAAATACCAATCATATC 868

910 CCGGGCAACAGAGTGGAGCCCTGGTC'AAAGAAAAA'AAAAAGCAACATATC 969

486 CCTGGGCAACAGAGTGAGACCTGTCTAAAGAA-----AAAAAAATAAAGGAACTATC 540

486 C C T G G G C A C A G A G T G A G A C C C T G T C T A A A G A A - - - - - A A A A A A A A T A A A G C A A C A T A T C 540

970 CTAAATAAAGGATCCTCCATAATGTTTCCACCAGATTTCATAATCAGAAACATGGAGGCCA 1029

370 C T A A A T A G A A G G A T C C T C C A T A A T G T T T C C A C C A G A T T C T A A T C A G A A T A C A T G G A G G C C A T G 22

Db	541	CTGAACAAAGGATCCTCATAAAGTTCCTCCACCAAGATTTCATAACAGAAAACATGGAGGCCA	600
Qy	1030	GGAAAGCAGTGGAGAAATGACGACACCTCAGGCAGAGCCCTGGAGGATGCTGTCAAGGCTGGG	1089
Db	601	GAAGCAGTGGAGGAGCGACCTCAGGCAGAGCCCGGAGGATGTTGTACAGGCTGGG	660
Qy	1090	CAAGGGCTTTCAAGCTACCAACTCTGGAGCTCTGGGAACAGCCCTGTTGCAACAGGAAGT	1149
Db	661	CAAGGGCTTTCCGCTACCAACTCTGGAGCTCTGGGAACAGCCCTGTTGCAACAGGAAGC	720
Qy	1150	CATGGCCCGGCAGAGCCCA-GAATGTGGCTGAGCTGGGATCCATGTGACAGCTTTGAG	1208
Db	721	CATAGCCCGGCCAGAGCCCAAGGAATGTGGGCTGGGCT-----	757
Qy	1209	GCTCACCGGAGCAGCTCTGGACAGGAGGTCCCATCCAGGAAACCTCGGCGATGGCT	1268
Db	758	-----GGAGCAGGCTCTGGACAGGAGTGGTCCATCCAGGAAACCTCCGCGATGGCT	810
Qy	1269	GGAAAGTGGGTTACTTTGGTCCGGGCTCTGATGTGTGTGACTGGTGTGTGAGAGAG	1328
Db	811	GGCAAGTGGGTTACTTTGGTCCGGGCTGTATGTGTGTGTGACTGGTGTGTGAGAGAG	870
Qy	1329	AATGTGCCCCCTCAGTGTCAAGTGTGAGTCTGTGTATGTGTGAATATTTCTTTGTGTGGG	1388
Db	871	AATGTGTGCCCTAAGTGTCAAGTGTGAGTCTGTGTATGTGTGAATATTTCTTTGTGTGGG	930
Qy	1389	TGATTTTCTGCATGTGTAATCTGTGCCCTGCAAGTGTGAACAAGTGACAAAGTGTCTGGG	1448
Db	931	TGATTTTCTCGTGTGTGTAATCTGTGTCCCTGCAAGTGTGAACAAGTGTGCAAGTGTCTGGG	990
Qy	1449	AGTGGACAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTGTGCAATGCAAGAGT	1508
Db	991	AGTGGACAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTGTGCAATGCAAGAGT	1050
Qy	1509	GCAAGGTGAAGTGAAGGACAGGCCCATGTGCCCATCATCATCAGGAGCTCTAAGGCC	1568
Db	1051	GCAAGGTGAAGTGAAGGACAGGCCCATGTGCCCATCATCATCAGGAGCTCTAAGGCC	1110
Qy	1569	CCAGGTAAGTGCAGGTGACAGATTAAGGGTCTCAAGGTCACTCTGGAGTGGGAGGTGGG	1628
Db	1111	CCAGGTAAGTGCAGGTGACAGATTAAGGGTCTCAAGGTCACTCTGGAGTGGGAGGTGGG	1170
Qy	1629	GGTAGGAAAGGCCAAGGTCAATGTTCTGGAGGAGGGTGTGTGACTACATPAGGGTGTATG	1688
Db	1171	GGTAGGAAAGGCCAAGGCATGTTCTGGAGGAGGGTGTGTGACTACATPAGGGTGTATG	1230
Qy	1689	AGCCTAGCTGGAGGTGGATGGCCGGGTCCACTGAGACCCCTGGTTATCCAGAAAGCCTGT	1748
Db	1231	AGCCTAGCTGGAGGTGGATGGCCGGGTCCACTGAAACCCCTGGTTATCCAGAAAGCCTTT	1290
Qy	1749	GTGGGCTTGGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCCGACCAAGCCCTTCTTAC	1808
Db	1291	GCAGGCTTCAGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCCGACCAAGCCCTTCTTAC	1350
Qy	1809	CACCTACCTGGGTAAAGGCCCTGGAGCAGGAAGCAGCGCAAGGACCTCTGGAGCAGCC	1868
Db	1351	CGGCTTACCTTGGTTAAAGGGCTTGGAGCAGGAAGCAGGGGCAAGAACCTCTGGAGCAGCC	1410
Qy	1869	CATACCTGCCCCTGACTCTGCCACTGGCAGCACAGTCAACACAGCAGAGTTCACTCA	1928
Db	1411	CATACCGGCCCTGGCCTGACTCTGCCACTGGCAGCACAGTCAACACAGCAGAGTTCACTCA	1470
Qy	1929	CAGCAGGGCGAAGGCCCATCATCAGCTCCCTTTATAAGGGAAGGCTCAGCGCCTGGTG	1988
Db	1471	CAGCAGGGCGAAGGCCCATCATCAGCTCCCTTTATAAGGGAAGGCTCAGCGCCTGGTG	1530
Qy	1989	TGCCGAGAGTGTCTGCCCTGGTCTCTGTGCCCTGGTGGGCTGGGGGTGCCAGGTGTGTCC	2048
Db	1531	TGCTGAGAGTGTCTGCCCTGGTCTCTGTGCCCTGGTGGGCTGGGGGTGCCAGGTGTGTCC	1590
Qy	2049	AGAGGAGCCCAAGTTGTTAGTGAGGACCATGGGGCTAGAAAGCACTGGTGTCCCTGGGCCA	2108


```
Db 5949 TCCTCAGGGCAACCCACCCCTGCTCAATTCCTGTTATCCCAACGCGCTGGCGCATTTG 6008
QY 6435 AGAC--GGGTAGCTGAGGCTGAGCAGATGTGATACCTTGCCCAATAATCCCATGTC 6492
Db 6009 AGACAGGGGTAGCTGAGGCTGAGCAGATGTGATACCTTGCCCAATAATCCCATGTC 6068
QY 6493 CCCACTGACCCCACTCTGACTGCCAGATTTGTGACAGGACTACATGTTCTTGGCATGT 6552
Db 6069 CCCACTGACCCCACTCTGACTGCCAGATTTGTGACAGGACTACATGTTCTTGGCATGT 6128
QY 6553 GGGGAAGGGCCAGAAATGGCTGACTAGAGTGTGATGAGTGTGATGAGTGTGATGAG 6612
Db 6129 GGGGAAGGGCCAGAAATGGCTGACTAGAGTGTGATGAGTGTGATGAGTGTGATGAG 6188
QY 6613 GGCAGGACTCAGCTGGAGGCCATATTTTCAGGCTTAATCAGCCACCCCAATCAGG 6672
Db 6189 GGCAGGACTCAGCTGGAGGCCATATTTTCAGGCTTAATCAGCCACCCCAATCAGG 6248
QY 6673 ACAGCAGTCTGCGCAGCACCACACAGTCACTCCCTTCATATGACACCCCAAAA 6732
Db 6249 ACAGCAGTCTGCGCAGCACCACACAGTCACTCCCTTCATATGACACCCCAAAA 6308
QY 6733 TGGAGACAAATCAT---GTCAGGAGCTATATGCCAGGG----CTACCTCCCGAGGCTC 6785
Db 6309 CGGAGACAAATCATGGCTCAGGAGCTATATGCCAGGCTACCTACCTCCCGAGGCTC 6368
QY 6786 AGTCGCGAGTGCAGAACTTCCCTGGGAAGGCCCGAGGAAACCCAGGACCGAGCCAC 6845
Db 6369 AGTCGCGAGTGCAGAACTTCCCTGGGAAGGCCCGAGGAAACCCAGGACCGAGCCAC 6428
QY 6846 CGCCCTCAGCCT 6857
Db 6429 CACCTCAGCCT 6440
```

RESULT 4

```
ACA61302
ID ACA61302 standard; DNA; 9433 BP.
XX ACA61302;
XX
XX 16-JUL-2003 (first entry)
XX Human cytochrome p450 gene CYP2D6, variant sequence.
XX
XX Human; ds; gene; cytochrome P450; CYP2D6; chromosome 22; SNP;
KW single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
KW psychiatric disorder; drug sensitivity.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(226..227,ATT)
FT /tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace(971,G)
FT /tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1111,T)
FT /tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1726,C)
FT /tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1846,A)
FT /tag= f
FT /standard_name= "Single nucleotide polymorphism"
FT replace(1846,G)
FT /tag= e
FT /standard_name= "Single nucleotide polymorphism"
FT replace(2064,A)
FT /tag= g
FT /standard_name= "Single nucleotide polymorphism"
FT
```

```
FT variation replace(3023,A)
FT /tag= h
FT /standard_name= "Single nucleotide polymorphism"
FT replace(5799,G)
FT /tag= i
FT /standard_name= "Single nucleotide polymorphism"
FT replace(5816..5817,C)
FT /tag= j
FT /standard_name= "Single nucleotide polymorphism"
XX
PN EPI281755-A2.
XX
XX 05-FEB-2003.
XX
XX 16-JUL-2002; 2002EP-00254972.
XX
XX 31-JUL-2001; 2001US-0309111P.
XX (PFIZ ) PFIZER PROD INC.
XX
XX Milos PM, Webb SM;
XX
XX WPI; 2003-373769/36.
XX
XX New cytochrome P450 2D6 gene variants and polypeptides, useful for
XX determining if a subject has or is at risk of developing a drug
XX sensitivity condition or disorder that is associated with an aberrant
XX CYP2D6 activity.
XX
XX Claim 3; Fig 3; 88pp; English.
```

The invention relates to an isolated nucleic acid comprising a cytochrome P450 2D6 gene variant, e.g. G5799C or C5816AT (referring to the genomic sequence or the same variant nucleotide in the corresponding cDNA sequences). Also included are probes, primers (allele specific oligonucleotides) and arrays used to detect and/or amplify the CYP2D6 gene polymorphic regions, the variant polypeptides, antibodies which are capable of distinguishing between the variant and wild-type polypeptides, determining whether a subject has a genetic deficiency for metabolising a drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and determining whether an individual is susceptible to being a poor metaboliser of drugs. The DNA probe is useful for hybridising to a variant form of the CYP2D6 gene. The primer is useful for amplifying the C5816TA allelic variant. The allele specific nucleotide is useful for the detection of the C5816TA allelic variant. The methods are useful for determining whether a subject has a genetic deficiency for metabolising a drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and determining if an individual is susceptible to being a poor metaboliser of drugs. The nucleic acids are useful as probes or primers for determining whether a subject has a genetic deficiency for metabolising drugs that are substrates of P450 CYP2D6. The methods are useful for determining if a subject has or is at risk of developing a drug sensitivity condition or disorder that is associated with an aberrant CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an aberrant CYP2D6 bioactivity. The methods are also useful in selecting the appropriate drugs or determining the course of treatment to administer to a subject to treat cardiovascular or psychiatric disorders, or for treating a subject with a drug sensitivity or disorder associated with a specific allelic variant of a polymorphic region of the CYP2D6 gene. The antibodies are useful for monitoring CYP2D6 protein levels in an individual for determining whether a subject has a disease or conditions associated with an aberrant CYP2D6 protein level. The gene is located on human chromosome 22. The present sequence is the variant CYP2D6 gene carrying both the G5799C and C5816AT variations

SQ Sequence 9433 BP; 1965 A; 2647 C; 2975 G; 1846 T; 0 U; 0 Other;

Query Match 51.7%; Score 5314.6; DB 7; Length 9433;

Best Local Similarity 93.4%; Pred. No. 0;

Matches 5895; Conservative 0; Mismatches 269; Indels 149; Gaps 27;

QY 637 TGTCAAAAAATCAGCCAGGATGTTGGTGGGCACTTATCCAGTACTTGGAGGC 696

Db 186 TGAATAACAAAAGCTAGAGCTGGTGGGCACACACCTGTAAATCCAGCTACTTAGAGGC 245
Qy 697 TGAGGCAGGAGAAATCGCTTGAACCTCGAGGTGGAGGTTCAGGTGAGCCGAGATCGCATC 756
Db 246 TGAGGCAGGAGAAATCGCTTGAACCTTAGAGGTGAAGGTTCAGGTGAGCCGAGATTCATC 305
Qy 757 ATTGCA-----CTCCACCTGGCGAC--AGAGCAAAATCTCGTCTC----- 796
Db 306 ATTGCAAAATGGAGGGAGCCAGCCTGGGCAACAAGAGGAATCTCCGTCTCAAAA 365
Qy 797 -----AAAAAATAATTTGGCTCGGTGGTGGCTGTAGTCCAGCTACTTGGGAG 849
Db 366 AAAAAAAAAAAAAAAAAAAGATTAGGCTGGTGGTGGCTGTAGTCCAGCTACTTGGGAG 425
Qy 850 GCAGGAGTCCACTGTGATGTGAGATTGCAATGAGTGCAGCCATGATCTGCCACTGCCG 909
Db 426 GCAGGGGTCCACTGTGATGTGAGACTGCAATGAGTGCAGTGCAGCCATGATCTGCCACTGCCG 485
Qy 910 CCGGGCAACAGAGTGCAGCCCTGTCTAAAGAAAAAATAAATAAAGCAACATATC 969
Db 486 CCGGGCAACAGAGTGCAGCCCTGTCTAAAGAA-----AAAAAATAAAGCAACATATC 540
Qy 970 CTAATAAAGGATCTCTCAATATGTTTCCACCAAGATTTCTAATCAGAAACATGAGGCCA 1029
Db 541 CTGACAAAGATCTCTCAATAGCTTCCACCAAGATTTCTAATCAGAAACATGAGGCCA 600
Qy 1030 GGAAGCAGTGGAGATGACGACCTCAGGCAGCCCTGGAGGATGCTGTCACAGCTGGG 1089
Db 601 GAAAGCAGTGGAGAGGAGCACCCTCAGGCAGCCCGGAGGATGTTGTACAGCTGGG 660
Qy 1090 CAAGGGCTTCAGGCTACCACTGGAGCTCTGGAAACAGCCCTGTTGCAAAACAGAAAT 1149
Db 661 CAAGGGCTTCAGGCTACCACTGGAGCTCTGGAAACAGCCCTGTTGCAAAACAGAAAT 720
Qy 1150 CATGGCCGCGCAGAGCCCA--GAATGTGGCTGAGCTGGGATCCATGTGACACTTGAG 1208
Db 721 CATAGCCGCGCAGAGCCCAAGGAATGTGGCTGGGCT----- 757
Qy 1209 GCTCAGCGGAGCAGCTCTGGACAGAGAGGTCCCATCCAGGAAACCTCGGCAATGGCT 1268
Db 758 -----GGGAGCAGCTCTGGACAGAGTGGTCCCATCCAGGAAACCTCGGCAATGGCT 810
Qy 1269 GGAAGTGGGATCTTGGTCCGGGCTGATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1328
Db 811 GGAAGTGGGATCTTGGTCCGGGCTGATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 870
Qy 1329 AATGTGTCCCTGAGTGTGAGT 1388
Db 871 AATGTGTCCCTAAGTGTGAGT 930
Qy 1389 TGAATTTCTGCATGTGTAATCGTGTCCCTGCAAGTGTGAACAGTGCACAGTGTCTGGG 1448
Db 931 TGAATTTCTGCATGTGTAATCGTGTCCCTGCAAGTGTGAACAGTGCACAGTGTCTGGG 990
Qy 1449 AGTGGACAGAGATCTGTGCACCATCAGTGTGTGCATAGCGTCTGTGCATGTCAAGAT 1508
Db 991 AGTGGACAGAGATCTGTGCACCATCAGTGTGTGCATAGCGTCTGTGCATGTCAAGAT 1050
Qy 1509 GCAAGTGAAGTGAAGGACAGGCCCATGATGCCATCATCATCAGGAGCTTAAGGCC 1568
Db 1051 GCAAGTGAAGTGAAGGACAGGCCCATGATGCCATCATCATCAGGAGCTTAAGGCC 1110
Qy 1569 CCAAGTGAAGTGCAGTGCACAGATAAGGGTGTCTGAAGTCACTCTGAGTGGGAGGTGG 1628
Db 1111 CCAAGTGAAGTGCAGTGCACAGATAAGGGTGTCTGAAGTCACTCTGAGTGGGAGGTGG 1170
Qy 1629 GGTAGGAAAGGCAAGGTCAATGTTCTGAGGAGGGGTTGTGACTACATTTAGGGTGTATG 1688
Db 1171 GGTAGGAAAGGCAAGGTCAATGTTCTGAGGAGGGGTTGTGACTACATTTAGGGTGTATG 1230
Qy 1689 AGCCTAGCTGGAGGTGGATGGCCGGGTCCATGACACCTGTGTTATCCAGAGCCTGT 1748
Db 1231 AGCCTAGCTGGAGGTGGATGGCCGGGTCCATGACACCTGTGTTATCCAGAGCCTTTT 1290

Qy 1749 GTGGCTTGGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCCGACCAAGCCCTTCTTAC 1808
Db 1291 GCAGGCTTTCAGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCCGACCAAGCCCTTCTTAC 1350
Qy 1809 CACCTTACCTTGGGTAAAGGGCTTGGAGCAGGAGAGAGCGGCAAGGACCTCTGAGAGAGCC 1868
Db 1351 CGGCTTACCTTGGGTAAAGGGCTTGGAGCAGGAGAGAGCGGCAAGGACCTCTGAGAGAGCC 1410
Qy 1869 CATACCTGCTTGGCTGACTCTGCCACTTGGCAGCAGTCAACACAGCAGGTTCACCTCA 1928
Db 1411 CATACCTGCTTGGCTGACTCTGCCACTTGGCAGCAGTCAACACAGCAGGTTCACCTCA 1470
Qy 1929 CAGCAGAGGCGGAAGCCCATCATCAGCTCCCTTTTATAAGGGAAGGGTTCAGCGCTCGGTG 1988
Db 1471 CAGCAGAGGCGGAAGCCCATCATCAGCTCCCTTTTATAAGGGAAGGGTTCAGCGCTCGGTG 1530
Qy 1989 TGGCAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2048
Db 1531 TGTGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1590
Qy 2049 AGAGGAGCCAGTGTGTAGTGGAGCAGCATGGGCTAGAGCAGTGTGCTGCTGCTGCTGCTG 2108
Db 1591 AGAGGAGCCCATTTGTGTAGTGGAGCAGGTATGGGCTAGAGCAGTGTGCTGCTGCTGCTGCTG 1650
Qy 2109 TGTAGTGGCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2168
Db 1651 TGTAGTGGCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1710
Qy 2169 CAGCTTACCGGCGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2228
Db 1711 CAGCTTACCGGCGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1769
Qy 2229 TTTCCAGAACACACCTACTGCTTCCGACAGGTGAGGAGGAGTCTCTGAGGCGGCGCAGA 2288
Db 1770 TTTCCAGAACACACCTACTGCTTCCGACAGGTGAGGAGGAGTCTCTGAGGCGGCGCAGA 1829
Qy 2289 GGTCTCTGAGGATGCCCCACACACAGCAACATGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2348
Db 1830 GGTCTCTGAGGCTCCCTTACCAGAACATGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1889
Qy 2349 CAGAGCCAGCTGAGAGAGGGAAGCAGGTCTTGGGGAAGTCTCTGAGGGAAGGACATTTA 2408
Db 1890 CAGAGCCAGCTGAGAGAGGGAAGCAGGTCTTGGGGAAGTCTCTGAGGGAAGGACATTTA 1948
Qy 2409 TACATGGCATGAAGGACTGGATTTTCCAAAGCCAAAGAGAGTAGGGCAAGGCGCTGGA 2468
Db 1949 TACATGGCATGAAGGACTGGATTTTCCAAAGCCAAAGAGAGTAGGGCAAGGCGCTGGA 2008
Qy 2469 GGTGAGCTGGACTTGGCAGTGGGATGCAAGCCATTTGGGCAACATATGTTATGAGTA 2528
Db 2009 GGTGAGCTGGACTTGGCAGTGGGATGCAAGCCATTTGGGCAACATATGTTATGAGTA 2068
Qy 2529 CAAAGTCCCTTCTGCTGACACAGAGGAAGGCTTGGGAATGGAAGATGAGTGTAGTTC 2588
Db 2069 CAAAGTCCCTTCTGCTGACACAGAGGAAGGCTTGGGAATGGAAGATGAGTGTAGTTC 2128
Qy 2589 TGAGTGGCTTTAAATCAGGAATCGAGATGAAAGGGGTGAGTGGGCTTCAAAAC 2648
Db 2129 TGAGTGGCTTTAAATCAGGAATCGAGATGAAAGGGGTGAGTGGGCTTCAAAAC 2188
Qy 2649 TTTTTCAGCTTGGGCTTCTCGGCTTCACTG-CTCACCGGATGAGGACCATCATCTGGGAAT 2707
Db 2189 TTTTTCAGCTTGGGCTTCTCGGCTTCACTGCTCACCGGATGAGGACCATCATCTGGGAAT 2248
Qy 2708 GGGATGCTAACTGGGCTTCTCGGCAATTTTGGTGAATTTTGGTGAATTTTGGTGAATTTTGGTGA 2767
Db 2249 GGGATGCTAACTGGGCTTCTCGGCAATTTTGGTGAATTTTGGTGAATTTTGGTGAATTTTGGTGA 2308
Qy 2768 AGGATCCAACTGAGTTCCTCCATCAGAGAGGTGTGACCCCGGCTGCTGCTGCTGCTGCTGCTGCTG 2827
Db 2309 AGGATCCAACTGAGTTCCTCCATCAGAGAGGTGTGACCCCGGCTGCTGCTGCTGCTGCTGCTGCTG 2368

Qy	2828	CAGAGGCTGGGTCTCTCTCTTCCACTGTCTCACTCTCTGTAGCCCGGGGTGTCCTCAA	2887
Db	2369	CAGAGGCTGGGTCTCTCTCTTCCACTGTCTCACTCTCTGTAGCCCGGGGTGTCCTCAA	2428
Qy	2888	GGTTCAAATAGGACTAGGACCTGTAGTCTGGGGTGATCTTGGCTTGACAAGAGGCCCTGA	2947
Db	2429	GGTTCAAATAGGACTAGGACCTGTAGTCTGGGGTGATCTTGGCTTGACAAGAGGCCCTGA	2488
Qy	2948	CCCTCCCTCTGCACTTCCGGCGCGCTTCCGGGACGTGTTTCAGCCTGACACTGGCCCTGGA	3007
Db	2489	CCCTCCCTCTGCACTTCCGGCGCGCTTCCGGGACGTGTTTCAGCCTGACACTGGCCCTGGA	2548
Qy	3008	CGCCGCTGGTCTGTCTCAATGGGCTGCGCCCTGTGCGAGGGCGATGGTGACCCCGCGCG	3067
Db	2549	CGCCGCTGGTCTGTCTCAATGGGCTGCGCCCTGTGCGAGGGCGCTGGTGACCCCGCGCG	2608
Qy	3068	AGGACA CGGCGGACCGCCGCTGCGCCCATCTACAGGTTCCTGGGCTTCGGGCGCGGTT	3127
Db	2609	AGGACA CGGCGGACCGCCGCTGCGCCCATCTACAGGTTCCTGGGCTTCGGGCGCGGTT	2668
Qy	3128	CCCAAGGCAAGCGCGCTGGGGACAGACCGCGTTTCCGTGGGCCCCGGGTGGACAGT	3187
Db	2669	CCCAAGGCAAGCGCGCTGGGGACAGACAG-ATTTCGTGGGACCGCGGTGGGTGAT	2726
Qy	3188	GACCGTAGCCAAAGACGCGCGACAGGGCTGGGGTCTTGGACGTGAACACAGAGATAAAG	3247
Db	2727	GACCGTAGTCCGAGCTGGGACAGAGGGCGCGGCTCGTGACATGAAC-----AG	2778
Qy	3248	GCCAGCCAGTGGGCTAGGACAGTGGGCGCAGGAAACCACTGCACGGGGAGGTGCGAGT	3307
Db	2779	GCCAGCCAGTGG-----GSA CAGCGGGCCAGAGAAACCACTGCATGAGGAGGTGAGC	2833
Qy	3308	CTGTGGGCTTG-----CAGGGGCGGGCTACTCCAGACCCCGCAGAA	3352
Db	2834	ATGGGACAGAGGGCGGGCTTTGTGA CAGTGTGGCGGGCCACTGCCAGACCTTGGCAGGA	2893
Qy	3353	GCCCGTGGGCGAGGCTG-----ATGCGTCCAAATGGCGGTGGCGG	3394
Db	2894	GCCCAATGGGTGAGCGTGGCGCATTTCCAGCTGGAATCCGGTGTGGAAGTGGGGCGGG	2953
Qy	3395	GACCGCCCTATGCTGGGGCTCATGTGTG-----GCGGACGGGCGGGATCTTCTTGA	3449
Db	2954	GACCGCACCTGTGCTGAAGCTCAGTGTGGGTGGCGGGGCCCGGGGCTCTTCCCTGA	3013
Qy	3450	GTGAAAGGTGTT CAGGCTGGGACAGACAGAGTGGGGCAAAACCCCGCCACAGGAGGG	3509
Db	3014	GTGCAAGGCGGT CAGGGTGGGACAGACAGAGTGGGG-CAAGACCTGCCCAAGCAAGG	3072
Qy	3510	GAGCAATGTGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGGCTAGGAGACTGC	3569
Db	3073	GAGCAAGTGGATGCA CAAAGATGGGCCCTGTGACCACTGGACAGAGCCAGGGACTGC	3132
Qy	3570	GGGAGACTTGTGGAGGCCCAGGGTTGGAGTGGGTGGCGAGGGTGGGCAAGGCCCTTC	3629
Db	3133	GGGAGACCAGGGGGAGC-ATAGGGTTGGAGTGGGTGGTGGATGGGGCTAATGCCCTTC	3191
Qy	3630	ATGSCAACGCCACGTCTCGTCCGCCGCCACAGGGGTATCTGTGCGCTATGGGCGCG	3689
Db	3192	ATGGCCACGGCACGTGCCCGTCCACCCCAAGGGGTGTTCTTGGCGCGCTATGGGCGCG	3251
Qy	3690	CGTGGCGGACGAGGCGCTTCTCGTGTCCACTTGGCGCAACTTTGGGCTTCGGGCAAGA	3749
Db	3252	CGTGGCGGACGAGGCGCTTCTCGTGTCCACTTGTGCCCACTTTGGGCTTCGGGCAAGA	3311
Qy	3750	AGTCGCTGGAGCAGTGGGTGACCGAGAGGCGCGCTGCTTTGTGCG-CTTCCGCGACC	3808
Db	3312	AGTCGCTGGAGCAGTGGGTGACCGAGAGGCGCGCTGCTTTGTGCGCGCTTCGCCCAACC	3371
Qy	3809	AAGCCGCTGGTGTGATGGGCAAGGGGCACAAAGGGGAACTGGGAGGGGGGACGGAG	3868
Db	3372	ACTCCGCTGGTGTGATGGGCAAGAGGGGCACAAAGCGGGAACTGGGAGGGGGGGACGGGG	3431
Qy	3869	AAGGCAACCCCTTACCGCGCATCTCCCCACCCCGGAGGACGCCCTTTTCCGCCCAACGCCCT	3928

[illegible]

FT	variation	as K in the specification"	1468	CACCATCAGGTGTGTGATAGCGTCTGTGATGTCAAGAGTGAAGGTGAAGGGA	1527
FT		replace(2118, T)			
FT		/*tag= ac			
FT		/label= PS24			
FT		/note= "Novel single nucleotide polymorphism (SNP); given			
FT	variation	as Y in the specification"	391	CACCATCAGGTGTGTGATAGCGTCTGTGATGTCAAGAGTGAAGGTGAAGGGA	450
FT		replace(2170, A)			
FT		/*tag= ad			
FT		/label= PS25			
FT		/note= "Known single nucleotide polymorphism (SNP); given			
FT	variation	as R in the specification"	1528	CCAGGCCCATGATGCCACTCATCATCAGGAGCTCTTAAGGCCCCAGGTAAAGTGAAGGGA	1587
FT		replace(2179, C)			
FT		/*tag= ae			
FT		/label= PS26			
FT		/note= "Novel single nucleotide polymorphism (SNP); given			
FT	exon	as S in the specification"	451	CCAGGCCCATGATGCCACTCATCATCAGGAGCTCTTAAGGCCCCAGGTAAAGTGAAGGGA	510
FT		replace(2179, C)			
FT		/*tag= af			
FT		/number= 3			
FT	variation	replace(2611, A)	1588	AGATAAGGCTGTGAAGGTCACTCTCGAGTGGGAGTGGGGTGGGGTGGGGAGGCAAGGT	1647
FT		/*tag= ag			
FT		/label= PS27			
FT		/note= "Novel single nucleotide polymorphism (SNP); given			
FT	variation	as W in the specification; causes the amino acid	511	AGATAAGGCTGTGAAGGTCACTCTCGAGTGGGAGTGGGGTGGGGTGGGGAGGCAAGGT	570
FT		substitution F120I"			
FT		replace(2635, C)			
FT		/*tag= ah			
FT		/label= PS28			
FT		/note= "Novel single nucleotide polymorphism (SNP); given			
FT	variation	as Y in the specification; causes the amino acid	1648	CATGTTCTCGAGAGGGGTTGTGACTACATTAGGTTGATGAGCCTAGCTGGGAGGTGGA	1707
FT		substitution W128R"			
FT		replace(2659, A)			
FT		/*tag= ai			
FT		/label= PS29			
FT		/note= "Novel single nucleotide polymorphism (SNP); given			
FT	variation	as R in the specification; together with PS30 causes the	571	CATGTTCTCGAGAGGGGTTGTGACTACATTAGGTTGATGAGCCTAGCTGGGAGGTGGA	630
FT		amino acid substitution V136I"			
FT		replace(2661, C)			
FT		/*tag= aj			
FT		/label= PS30			
FT		/note= "Known single nucleotide polymorphism (SNP); given			
FT	Query Match	49.3%; Score 5070.2; DB 6; Length 6472;	1708	TGGCCGGGTCCACTGAGACCCCTGGTTATCCAGAGCCCTGTGTGGGCTTTGGGAGGCTTGG	1767
FT	Best Local Similarity	93.8%; Pred. No. 0;			
FT	Matches 5507; Conservative	41; Mismatches 209; Indels 116; Gaps 22;	631	TGGCCRRGGTCCACTGAAACCCCTGGTTATCCAGAGGGCTTTGCAGGCTTCAGGAGCTTGG	690
Qy	1049	GACCCCTCAGGAGCCCTGGAGGATGCTGTACAGGCTGGGGCAAGGGCCTTCAGGCTACC	1768	AGTGGGAGAGGGGTGACTTCTCCGACCAAGGCTTTTCTACCACTTACCTCCCTGGGTAAAGG	1827
Db	1	GACCCCTCAGGAGCCCTGGAGGATGCTGTACAGGCTGGGGCAAGGGCCTTCAGGCTACC			
Qy	1109	AACCTGGGAGCTCTGGGAACAGCCCTGTGCAAAACAGGAAGTCAATGGCCCGCCAGAGCCC	691	AGTGGGAGAGGGGTGACTTCTCCGACCAAGGCTTTTCTACCACTTACCTCCCTGGGTAAAGG	750
Db	61	AACCTGGGAGCTCTGGGAACAGCCCTGTGCAAAACAGGAAGTCAATGGCCCGCCAGAGCCC			
Qy	1169	A-GAATGTGGGCTGAGCTGGGATCCATGTGACAGCTTTTGAGGCTCACCGGAGCAGCCTC	1828	GCCTGGAGCAGGAAGCAGCGGCAAGGACCTCTGGAGCAGGCCCATACCTGCTCCCTGGGCTGA	1887
Db	121	AGGAATGTGGGCTGGGCT-----GGGAGCAGCCTC			
Qy	1228	TGACAGGAGAGTCCCATCAGGAACCTCGGGCATGGCTGGGAAGTGGGGTACTTGGT	751	GCCTGGAGCAGGAAGCAGCGGCAAGGACCTCTGGAGCAGGCCCATACCTCCCTGGGCTGA	810
Db	151	TGACAGGAGTGGTCCCATCAGGAACCTCCGGCATGGCTGGGAAGTGGGGTACTTGGT			
Qy	1288	GCCGGGTCTGTATGTGTGTGACTGGTGTGTGTGAGAGAGATGTGTGCCCTGAGTGTC	1888	CTCTGCCACTTGGCAGCAGTCAACACAGAGAGTTTCACTCAAGCAGAGGGCCGGAAGGCCA	1947
Db	211	GCCGGGTCTGTATGTGTGTGACTGGTGTGTGTGAGAGAGATGTGTGCCCTGAGTGTC			
Qy	1348	AGTGTGAGTCTGTATGTGTGATATTTCTTTGTGGGTGATTTCTTGTGATGTGTA	811	CTCTGCCACTTGGCAGCAGTCAACACAGAGAGTTTCACTCAAGCAGAGGGCCGGAAGGCCA	870
Db	271	AGTGTGAGTCTGTATGTGTGATATTTCTTTGTGGGTGATTTCTTGTGCTGTGTA			
Qy	1408	TCGTGCTCCCTGCAAGTGTGAACAGTGGCAAGTGTCTGGAGTGGCAAGAGATCTGTG	1948	TCATCAGCTCCCTTTTATAGGGAAGGTCACCGCTCGGTGTCGCCAGAGAGTGTCCCTGCCT	2007
Db	331	TCGTGCTCCCTGCAAGTGTGAACAGTGGCAAGTGTCTGGAGTGGCAAGAGATCTGTG			
			871	TCATCAGCTCCCTTTTATAGGGAAGGTCACCGCTCGGTGTCGCCAGAGAGTGTCCCTGCCT	930
			2008	GCTCCTCTGTGCTGGTGGGGTGGGGGTGCCAGGTGTGTCAGAGGAGCCAGTGTGGTAG	2067
			931	GCTCCTCTGTGCTGGTGGGGTGGGGGTGCCAGGTGTGTCAGAGGAGCCCAATTTGGTAG	990
			2068	TGAGGAGCAGCATGGGCTAGAGCACTGGTGGCCCTTGGCCATGATAGTGCCCATCTTCTCT	2127
			991	TGAGGAGCAGTATGGGCTAGAGCACTGRTGGCCCTTGGCCRTGATAGTGCCCATCTTCTCT	1050
			2128	GCTCCTGTGTGACCTGATGCACCGGCAACCAAGCTGGGCTGACGCTACCCGCGAGTCC	2187
			1051	GCTCCTGTGTGACCTGATGCACCGGCGCAAGCTGGGCTGACACGCTTCTCAACAGGCC	1110
			2188	CCTGCCACTGCCCGGCTGGGCAACCTTGTGTCATGTGGACTTTCAGAAACACACCATACT	2247
			1111	CCTGCCACTGCCCGGCTGGGCAACCTTGTGTCATGTGGACTTTCAGAAACACACCATACT	1169
			2248	GCTTCCAGCAGGTGAGGAGAGGAGTCTCGAGGCGGCGAGAGTCTCTGAGGATGCCCCAC	2307
			1170	GCTTCCAGCAGGTGAGGAGAGGAGTCTCGAGGCGGCGAGAGTCTCTGAGGCTCCCTTAC	1229
			2308	CACCAAGCAACATGGGTGGGTAAACCAACAGGCTGGATCAGAGAGCCAGGCTGAGAAG	2367
			1230	CAGAAGCAACATGGATGGGTGAAACCAACAGGCTGGACCAAGAGCCAGGCTGAGAAG	1289
			2368	GGGAACAGGTTTGGGGGACGTTTCTCGGGGAGGACATTTATACATGGCATGAAGACTG	2427
			1290	GGGAACAGGTTTGGGGGACGTTTCTCGGGGAGGACATTTATACATGGCATGAAGACTG	1348
			2428	GATTTTCCAAAGCCCAAGGAAGTAGAGGCAAGGGCTCGAGGTGGAGCTGGACTTGGCA	2487
			1349	GATTTTCCAAAGCCCAAGGAAGTAGAGGCAAGGGCTCGAGGTGGAGCTGGACTTGGCA	1408
			2488	GTGGGATGCAAGCCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGAC	2547
			1409	GTGGGATGCAAGCCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGAC	1468

Qy	2548	ACCAAGAGAAAGCCCTTGGGAATGGAAGATGATGATCTCTGAGTGCCCTTTAAATCAC	2607
Db	1469	ACCAGAAGGAAAGCCCTTGGGAATGGAAGATGATGATCTGAGTGCCCGTTTAAATCAC	1528
Qy	2608	GAAATCGAGATGAAGGGGTGCAGTGACCCGGTTCAAACTTTTCACACTGTGGGTCCCTC	2667
Db	1529	GAAATCGAGATGAAGGGGTGCAGTGACCCGGTTCAAACTTTTCACACTGTGGGTCCCTC	1588
Qy	2668	GGGCCTCACTG-CTCACCCGCATGGACCATCATCTCGGAATGGGATGCTAACTGGGGCCT	2726
Db	1589	GGGCCTCACTGCTCACCCGCATGGACCATCATCTCGGAATGGGATGCTAACTGGGGCCT	1648
Qy	2727	CTGGCAATTTTGGTGACTCTTGAAAGGTCATACCTGGGTGACGCATCCAAACTGAGTTC	2786
Db	1649	CTGGCAATTTTGGTGACTCTTGCAAGGTCAATACCTGGGTGACGCATCCAAACTGAGTTC	1708
Qy	2787	CTCCATCACAGAAGGTGTGACCCCAACCCCTGCCCCACGATCAGGAGGCTGGGTCTCCCTC	2846
Db	1709	CTCCATCACAGAAGGTGTGACCCCAACCCCGCCGCCACGATCAGGAGGCTGGGTCTCCCTC	1768
Qy	2847	CTTCCACTGCTCACTCCTGCTAGCCCGGGGTCTGTCAAAGTTCAAAATAGGACTAGGA	2906
Db	1769	CTTCCACTGCTCACTCCTGCTAGCCCGGGGTCTGTCAAAGTTCAAAATAGGACTAGGA	1828
Qy	2907	CCTGTAGTCTGGGTGATCTCTGGCTTGACAAGAGGCCCTGACCCCTCCCTCTGCAGTTGCG	2966
Db	1829	CCTGTAGTCTGGGKGATCTCTGGCTTGACAAGAGGCCCTGACCCCTCCCTCTGCAGTTGCG	1888
Qy	2967	GGCCCGCTTCGGGACGTGTTTACGCTGACGCTGGCTTGACGCGCGTGTCTGCTCAA	3026
Db	1889	GGCCCGCTTCGGGACGTGTTTACGCTGACGCTGGCTTGACGCGCGTGTCTGCTCAA	1948
Qy	3027	TGGGCTGGCGCGTGGCGAGCGATGTGTACCCGCGCGAGGACACGCGCCACCGCCC	3086
Db	1949	TGGGCTGGCGCGTGTGRCAGGCGMTGTGTACCCRCGCGAGGACACGCGCCACCGCCC	2008
Qy	3087	GCCTGCCCCATCTACAGGTCCTGGGCTTTCGGCGCGCTTCCAAAGGCAAGCGCGGTG	3146
Db	2009	GCCTGYCCCATCWYCCAGRTCTCTGGGTYTYYGGCGCGTTCCTCAAGGCAAGCRGCGG-K	2067
Qy	3147	GGGGACAGACCGCGTTTCCGTGGGCCCCGGGTGTGACAGTGCAGTGCAGCCCAAGCAGCG	3206
Db	2068	GGGGACAGACAG-ATTTCCGTGGGACCCCGGGTGGGTGATGACCGTAGTCYAGCTGGG	2126
Qy	3207	CCGACAGGCGTGGGGTCTGTGCAGCTGAAACAGAGATAAAGGCGCAGCGATGGGCTGAGG	3266
Db	2127	CAGNAGGCGCGGGGTCTGTGGACATGAAC-----AGGCCAGGAGTGG-----GG	2173
Qy	3267	ACAGTGGGCGAGAAACACCTGCACGGGGAGGTGCGAGTCTGTGGGCTGG-----	3318
Db	2174	ACAGCGGCCCAAGAAACACCTGCACTAGGGAGGTGTGACATGGGACAGAGGCGGGGC	2233
Qy	3319	-----GAGGGCGGGGCTACTGCCACAGCCGCCAGAAAGCCCGGTGGGCGAGGCTG-	3370
Db	2234	TTGTGACAGTGGGCGGGGCACCTGCCGAGACCTTGGCAGGAGCCCAATGGGTGAGCGTGG	2293
Qy	3371	-----ATCGCTCGAGTGCAGTGGCGTGGCGGGGACCGCGCTATGCTGCGG	3413
Db	2294	CGCATTTCCAGCTGGGAATCCGTTGTGCAAGTGGGGCGGGGACCGCACCTGTGCTGTAA	2353
Qy	3414	GCTCAGTGTGGGCGGACGG-----GCGGGATCTTCTTGGTGGAAAGTGTGTACGGGT	3468
Db	2354	GCTCAGTGTGGGTGGCGCGGGGCCCGCGGGGTCTTCCCTGAGTGCAAAGCGCGTCAAGGT	2413
Qy	3469	GGGCAGACGAGGTGGGGCAAAACCCCGCCCCACAGGACGGGGAGCAATGTGGGTGAGCAA	3528
Db	2414	GGGCAGACGAGGTGGGG-CAAAAGCTGCTCCACGCAAGGAGCAAGGTGGATGCACAA	2472
Qy	3529	AGAGTGGGCCCTGTGCCACCTGACCGGGCTAGGGACTGCGGGAGACCTTGTGGAGCGC	3588
Db	2473	AGAGTGGGCCCTGTGACACACTGACACGACGAGGACTCTCGGAGACACAGGGGGAGC-A	2531
Qy	3589	CAGGGTTGAGTGGGTGGCGAGGGTGGGGCCCAAGGCTTTTCATGGCAACGCCACCTGTCT	3648

Db 3611 GGAGAGGTGAGTGGCTGCCACGGTGGGGGCAAGGGTGGTGGGTGAGCGTCCAGG 3670
Qy 4727 AGGAATGAGGGAGGTGGCAAAAGTTGGACAGTGCATACCCGGGAGCGGCATCT 4786
Db 3671 AGGAATGAGGGAGGTGGCAAAAGTTGGACAGTGCATACCCGGGAGCGCATCT 3730
Qy 4787 GGCGTACAGGTGCAGAAATGAGAGTCAATTTGGGGCTACCCGGTTCTATCCCTCAGTA 4846
Db 3731 GGCGTACAGGTGCAGAAATGAGAGTCAATTTGGGGGCTACCCGGTTCTGTGCC - GAGTA 3788
Qy 4847 TCCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCAATGATGAGAACCT 4906
Db 3789 TGCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCAATGATGAGAACCT 3848
Qy 4907 GCGCATAGTGGTGGGTAACTGTTCCTTGGCCGGATGGTGACCACTCGACACGCTGGC 4966
Db 3849 GCGCATAGTGGTGGCTGACCTGTTCCTGCGCGGATGGTGACCACTCGACACGCTGGC 3908
Qy 4967 CTGGGGCCTCTGCTCATGATCTACACCTGGATGTGACGCTGAGCCAGCTGGGGCCC 5026
Db 3909 CTGGGGCCTCTGCTCATGATCTTACATCCGATGTGACGCTGAGCCCATCTGGGAAAC 3968
Qy 5027 AAGCGAGGACTGAGGGAGGAAGGTACAGCTGGGGGCCCTCGGGCTTAGCTGGGACAC 5086
Db 3969 AGTGCGAGGGCCGAGGGAGGAGGTACAGGCGGGGCCCATGAATTTGCTGGGACAC 4028
Qy 5087 CGGGGCTTCAGACACAGCGCTGCGCAGGCTCTGTAAAGCCATACTTCTCCAAACACAGGA 5146
Db 4029 CGGGGCTCCAAACACAGGCTTGACAGGATCTGTAAAGCTGACCTCTCCAAACATAGGA 4088
Qy 5147 GGAAGGAGAGTGT-----CCCCCTGGGTGCTGACCCCATGTTGGGGGACGATGT 5193
Db 4089 GGAAGGAGAGTGTGAGGGCCGAGACCCCTGGGTGCTGACCCCATGTTGGGGGACGATGT 4148
Qy 5194 CTGTCCAGTCCGTCCACAGGAGATCGACAGCTGTAGGCGAGCTGGGGCGACAGA 5253
Db 4149 CTGTCCAGGCGGTGCCAAGGAGATCGACAGCTGTAGGCGAGGTGGCGGACAGGA 4208
Qy 5254 GATGGGTGACAGGCTCACATGCCCTACACCACTGCCGTGATTCACAGAGTGCAGCGCTT 5313
Db 4209 GATGGGTGACAGGCTCACATGCCCTACACCACTGCCGTGATTCAGAGTGCAGCGCTT 4268
Qy 5314 TGGGGACATCATCCCTGAGTGTGACCCCATATGACATCCGTGACATCGAAGTACAGGG 5373
Db 4269 TGGGGACATCGTCCCTCGGTGTGACCCCATATGACATCCGTGACATCGAAGTACAGGG 4328
Qy 5374 CTTCCGATCCCTAAGGTAGGCTGGGGCCCTCTCAGCCAGCTCAGCACGACCTG 5433
Db 4329 CTTCCGATCCCTAAGGTAGGCTGGGGCCCTCTCAGCCAGCTCAGCACGACCTG 4388
Qy 5434 GTGATAGCCCAAGCATGGCTACTGCGAGTGGGCCCACTCTAGGAACCTTGGCCACCTAG 5493
Db 4389 GTGATAGCCCAAGCATGGCTACTGCGAGTGGGCCCACTCTAGGAAMCCTTGGCCACCTAG 4448
Qy 5494 TCCTCAATGCCACACATGACTGTCCCACTTTGGGTGGGGGTCCAGAGTATAGGCAGG 5553
Db 4449 TCCTCAATGCCACACATGACTGTCCCACTTTGGGTGGGGGTCCAGAGTATAGGCAGG 4508
Qy 5554 GCTGGCCTGTCCATCAGAGCCCCGTCTAGTGGGAGACAAACAGGACCTGCCAGAAAT 5613
Db 4509 GCTGGCCTGTCCATCAGAGCCCCGTCTAGTGGGAGACAAACAGGACCTGCCAGAAAT 4568
Qy 5614 GTTGGAGGACCCAGCCCTGACGGAGAGGGGCGAGTGGGTGCTCTGAGAGGTGTGA 5673
Db 4569 GTTGGAGGACCCAAACCCCTGACGGAGAGGGGCGAGTGGGTGCTCTGAGAGGTGTGA 4628
Qy 5674 CTGCGCCCTGCTGTGGGGTCCGAGAGGGTACTGTGAGAGTCTCGGGCGCAGGACTAGTT 5733
Db 4629 CTGCGCCCTGCTGTGGGGTCCGAGAGGGTACTGTGAGAGTCTCGGGCGCAGGACTAGTT 4688
Qy 5734 GACAGAGTCAGTGTGTGCCAGGCAAGTGTGTGTCCCGGTGTGTGTGGCAGGGGTC 5793
Db 4689 GACAGAGTCAGTGTGTGCCAGGCAAGTGTGTGTCCCGGTGTGTGTGGCAGGGGTC 4748

Qy 5794 CCAGCATCTAGAGTCCAGTCCCACTCTCAACCTGCATCTCTCTGCCAGGGAACGACAC 5853
Db 4749 CCAGCATCTAGAGTCCAGTCCCACTCTCAACCTGCATCTCTCTGCCAGGGAACGACAC 4808
Qy 5854 TCATCACCAACCTGTCTCATCGGTGCTGAAGGATGAGCGCTCTGGGAGAGCCCTTCGGCT 5913
Db 4809 TCATCACCAACCTGTCTCATCGGTGCTGAAGGATGAGCGCTCTGGGAGAGCCCTTCGGCT 4868
Qy 5914 TCACCCCGAACAACCTTCCTGGATGCCAGGGCCACTTTGTGAAGCCGGAGGCCCTCTCTGC 5973
Db 4869 TCACCCCGAACAACCTTCCTGGATGCCAGGGCCACTTTGTGAAGCCGGAGGCCCTCTCTGC 4928
Qy 5974 CTTTCTCAGCAGGTGCTGTGGGAGCCCGGTCTCTGTGCCCTTCCGTGGAGTCTTGCA 6033
Db 4929 CTTTCTCAGCAGGTGCTGTGGGAGCCCGGTCTCTGTGCCCTTCCGTGGAGTCTTGCA 4988
Qy 6034 GGGGTATCACCCAGGAGCCAGGCTCATCTGACGCCCTTCCCTCCCAAGGCGCCGCTGC 6093
Db 4989 GGGGTATCACCCAGGAGCCAGGCTCATCTGACGCCCTTCCCTCCCAAGGCGCCGCTGC 5048
Qy 6094 ATGCTTCGGGGAGCCCTTGGCCCGCATGAGCTCTTCTCTTCTTTCACCTCCCTGCTGCA 6153
Db 5049 ATGCTTCGGGGAGCCCTTGGCCCGCATGAGCTCTTCTCTTCTTCTTCTTCTTCTTCTTCT 5108
Qy 6154 GCATTTAGCTTTCTCGTGGCCCGCGGACAGGCCCGGGCCCGGACCACTCTCTGTGTCTCAG 6213
Db 5109 GCATTTAGCTTTCTCGTGGCCCGCATGAGCACGCCCGGGCCCGGACCACTCTTCTTCTTCT 5168
Qy 6214 CTTTCTGTTGACCCCATCCCTAGAGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6273
Db 5169 TTTCTCTGT 5228
Qy 6274 TAGTCCCGAGCTGTCTCCCTAGCAGAGCTCTTAATGTACATAAAGCAATGTGTGTGTGT 6333
Db 5229 TAGTCCCGAGCTGTCTCCCTAGCAGAGCTCTTAATGTACATAAAGCAATGTGTGTGTGT 5288
Qy 6334 CCAACTTGGGTCTCCCTGTCTCAGCGCTCTCTTGGGATCATCTCTCCTCAGGGCAACCCAC 6393
Db 5289 CCAACTTGGGTCTCCCTGTCTCAGCGCTCTCTTGGGATCATCTCTCCTCAGGGCAACCCAC 5348
Qy 6394 CTGCTCATCTCTGCTTACCCACCGCTGCGCGCATTTGAGAC - GGGTACGTTGAGG 6451
Db 5349 CTGCTCATCTCTGCTTACCCACCGCTGCGCGCATTTGAGACAGAGGCTACGTTGAGG 5408
Qy 6452 CTGAGCAGATGTCTGCTTACCTTGGCCATAATCCCATGTCTCCCACTGACCCCACTCTGA 6511
Db 5409 CTGAGCAGATGTCTGCTTACCTTGGCCATAATCCCATGTCTCCCACTGACCCCACTCTGA 5468
Qy 6512 CTGCCCAGATTTGTCAGAGGACTACATTTCTCTGSCATGTGGGGAAGGGCGCAGAAATGG 6571
Db 5469 CTGCCCAGATTTGTCAGAGGACTACATTTCTCTGSCATGTGGGGAAGGGCGCAGAAATGG 5528
Qy 6572 GCTGATAGAGTGTCTGAGTCCCTGGATGTGGTGGAGGGGAGGACTCAGCCTGGAG 6631
Db 5529 GCTGATAGAGTGTCTGAGTCCCTGGATGTGGTGGAGGGGAGGACTCAGCCTGGAG 5588
Qy 6632 GCCCATATTTAGGCGCTTACTCAGCCCAACCCACATCAGGAGCAGAGTCTCTGCCAGCAC 6691
Db 5589 GCCCATATTTAGGCGCTTACTCAGCCCAACCCACATCAGGAGCAGAGTCTCTGCCAGCAC 5648
Qy 6692 CATCAACAGTCACTCTCCCTTATATATGACCCCAAAATGGGAAGACAAATCAT - -G 6748
Db 5649 CATCAACAGTCACTCTCCCTTATATATGACCCCAAAATGGGAAGACAAATCATGCGG 5708
Qy 6749 TCAGGGAGCTATATGCCAGG - - -CTACCTCCAGGGCTCAGTCCGCGAGGTGCCAGAAC 6804
Db 5709 TCAGGGAGCTATATGCCAGGCTACTTCTCCAGGGCTCAGTCCGCGAGGTGCCAGAAC 5768
Qy 6805 ATTCCCTGGGAAGGCCCGCAGGAAACCCAGAGCCGAGGCCACCGCCCTCAGCCT 6857
Db 5769 GTTCCCTGGGAAGGCCCGCAGGAGCCAGGACTGAGCCACCACTCAGCCT 5821


```

RESULT 6
AB072364
ID AB072364 standard; DNA; 6472 BP.
XX AC
XX AC
XX AB072364;
XX 02-SEP-2002 (first entry)
XX DE
XX Human CYP2D6 gene, SEQ ID NO:1 version #2.
XX
KW Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;
KW chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;
KW antiarrhythmic; arrhythmia; adrenoceptor antagonist; hypertension;
KW tricyclic antidepressant; procainamide; drug induced lupus syndrome;
KW environmentally linked disease; Parkinson's disease; haplotyping;
KW genotyping; haplotype; genetic variant; single nucleotide polymorphism;
KW SNP; drug screening; drug discovery; gene; ds.
XX OS
XX Homo sapiens.
XX
XX Key
XX Location/Qualifiers
XX FT variation
XX replace(636, A)
XX /tag= a
XX FT
XX /label= PS1
XX FT
XX /note= "Novel single nucleotide polymorphism (SNP)"
XX FT
XX replace(678, C)
XX /tag= b
XX FT
XX /label= PS2
XX FT
XX replace(769, C)
XX /tag= c
XX FT
XX /label= PS3
XX FT
XX /note= "Novel single nucleotide polymorphism (SNP)"
XX FT
XX replace(776, G)
XX /tag= d
XX FT
XX /label= PS4
XX FT
XX /note= "Novel single nucleotide polymorphism (SNP)"
XX FT
XX replace(825, A)
XX /tag= e
XX FT
XX /label= PS5
XX FT
XX /note= "Known single nucleotide polymorphism (SNP)"
XX FT
XX replace(915, C)
XX /tag= f
XX FT
XX /label= PS6
XX FT
XX /note= "Novel single nucleotide polymorphism (SNP)"
XX FT
XX 1001..5217
XX /tag= g
XX FT
XX /product= "CYP2D6"
XX FT
XX 1001..1180
XX /tag= h
XX FT
XX /number= 1
XX FT
XX replace(1019, A)
XX /tag= i
XX FT
XX /label= PS7
XX FT
XX /note= "Known single nucleotide polymorphism (SNP)"
XX FT
XX causes the amino acid substitution V7M"
XX replace(1031, A)
XX /tag= j
XX FT
XX /label= PS8
XX FT
XX /note= "Known single nucleotide polymorphism (SNP)"
XX FT
XX causes the amino acid substitution V11M"
XX replace(1100, T)
XX /tag= k
XX FT
XX /label= PS9
XX FT
XX /note= "Known single nucleotide polymorphism (SNP)"
XX FT
XX causes the amino acid substitution P34S"
XX 1181..1883
XX /tag= l
XX FT
XX /number= 1
XX FT
XX replace(1827, C)
XX /tag= m
XX FT
XX /label= PS10
XX FT

```

```

FT variation
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace(1843, G)
FT /tag= n
FT /label= PS11
FT /note= "Known single nucleotide polymorphism (SNP)"
FT 1884..2055
FT /tag= o
FT /number= 2
FT replace(1966, A)
FT /tag= p
FT /label= PS12
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT causes the amino acid substitution R88H"
FT replace(1974, A)
FT /tag= q
FT /label= PS13
FT /note= "Known single nucleotide polymorphism (SNP)"
FT causes the amino acid substitution L91M"
FT replace(1984, G)
FT /tag= r
FT /label= PS14
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT causes the amino acid substitution H94R"
FT replace(1997, G)
FT /tag= s
FT /label= PS15
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace(2014, C)
FT /tag= t
FT /label= PS16
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT causes the amino acid substitution V104A"
FT replace(2022, T)
FT /tag= u
FT /label= PS17
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT together with PS18 causes the amino acid substitution
FT T107F"
FT /tag= v
FT replace(2023, T)
FT /label= PS18
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT together with PS17 causes the amino acid substitution
FT T107F"
FT /tag= w
FT replace(2028, G)
FT /label= PS19
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT causes the amino acid substitution I109V"
FT replace(2036, C)
FT /tag= x
FT /label= PS20
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace(2039, T)
FT /tag= y
FT /label= PS21
FT /note= "Known single nucleotide polymorphism (SNP)"
FT 2056..2605
FT /tag= z
FT /number= 2
FT /cons_splice= (5'site:NO, 3'site:YES)
FT replace(2062, G)
FT /tag= aa
FT /label= PS22
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace(2067, G)
FT /tag= ab
FT /label= PS23
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace(2118, T)
FT /tag= ac
FT /label= PS24
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT

```


FT	variation		replace(2170, A)	
FT			/*tag= ad	
FT			/label= PS25	
FT			/note= "Known single nucleotide polymorphism (SNP) "	
FT	variation		replace(2179, C)	
FT			/*tag= ae	
FT			/label= PS26	
FT			/note= "Novel single nucleotide polymorphism (SNP) "	
FT	exon		2606. .2758	
FT			/*tag= af	
FT			/number= 3	
FT	variation		replace(2611, A)	
FT			/*tag= ag	
FT			/label= PS27	
FT			/note= "Novel single nucleotide polymorphism (SNP); causes the amino acid substitution F120I"	
FT	variation		replace(2635, C)	
FT			/*tag= ah	
FT			/label= PS28	
FT			/note= "Novel single nucleotide polymorphism (SNP); causes the amino acid substitution W128R"	
FT	variation		replace(2659, A)	
FT			/*tag= ai	
FT			/label= PS29	
FT			/note= "Novel single nucleotide polymorphism (SNP); together with PS30 causes the amino acid substitution V136I"	
FT	variation		replace(2661, C)	
FT			/*tag= aj	
FT			/label= PS30	
FT			/note= "Known single nucleotide polymorphism (SNP); together with PS29 causes the amino acid substitution V136I"	
FT	variation		replace(2704, G)	
FT			/*tag= ak	
FT			/label= PS31	
FT			/note= "Known single nucleotide polymorphism (SNP); causes the amino acid substitution Q151E"	
FT	variation		replace(2716, A)	
FT			/*tag= al	
FT			/label= PS32	
FT			/note= "Novel single nucleotide polymorphism (SNP); causes the amino acid substitution E155K"	
FT	intron		2759. .2846	
FT			/*tag= am	
FT			/number= 3	
FT	variation		replace(2846, A)	
FT			/*tag= an	
FT			/label= PS33	
FT			/note= "Known single nucleotide polymorphism (SNP)"	
FT	exon		2847. .3007	
FT			/*tag= ao	
FT			/number= 4	
FT	intron		3008. .3440	
FT			/*tag= ap	
FT			/number= 4	
FT	variation		replace(3292, A)	
Query Match		49.2%; Score 5061; DB 6; Length 6472;		
Best Local Similarity		94.2%; Pred. No. 0;		
Matches 5532; Conservative		0; Mismatches 225; Indels 116; Gaps 22;		
OY	1049	GACCCTCAGCGACGCCCTGGGAGGATGCTGCACAGGCTGGGGCAAGGCCCTTCAGGCTTACCC	1108	
Dd	1	GACCCTCAGCGACGCCCGGGAGGATTTGTTCACAGGCTGGGGCAAGGCCCTTCAGGCTTACCC	60	
OY	1109	AACTGGGAGCTTCGGGAACAGCCCTTGTTGCAAAACAGGAAGTCATGGCCCGGCCAGAGCCCC	1168	
Dd	61	AACTGGGAGCTTCGGGAACAGCCCTTGTTGCAAAACAGGAAGCCATAGCCCGGCCAGAGCCCC	120	
OY	1169	A-GAATGTGGGCTGAGCTGGGATCCATGTGACAGCTTTGAGGCTCATCCGGGAGCAGGCTC	1227	
Dd	121	AGGAATGTGGGCTGGGCT-----GGGAGCAGGCTC	150	

Qy	1228	TGACAGGAGAGTCCCATCCAGAAACCTTGGGCGATGGCTGGGAAGTGGGGTACTTTGGT	1287
Db	151	TGGACAGGAGTGGTCCCATCCAGAAACCTTCGGGCATGGCTGGGAAGTGGGGTACTTTGGT	210
Qy	1288	GCCGGGTCTGTATGTGTGTGTGACTGGTGTGTGAGAGAGAATGTGTGCCCTGAGTGTCT	1347
Db	211	GCCGGGTCTGTATGTGTGTGTGACTGGTGTGTGAGAGAGAATGTGTGCCCTTAAAGTGTCT	270
Qy	1348	AGTGTGAGTCTGTATGTGTGTAATATTTGCTTTGTGTGGGTGAATTTCTTGCATGTGTAA	1407
Db	271	AGTGTGAGTCTGTATGTGTGTAATATTTGCTTTGTGTGGGTGAATTTCTTCTCGCGTGTGTAA	330
Qy	1408	TCGTGTCCCTGCAAGTGTGTAAACAAGTGTGTGGAACAAGTGTCTGGGAATGTGAACAAGATCTGTG	1467
Db	331	TCGTGTCCCTGCAAGTGTGTAAACAAGTGTGTGGAACAAGTGTCTGGGAATGTGAACAAGATCTGTG	390
Qy	1468	CACCATCAGGTGTGCATAGCCGTCTGTGCATGTCAAGAGTGCAAAGTGAAGGTGAAGGA	1527
Db	391	CACCATCAGGTGTGTGATAGCGTCTGTGCATGTCAAGAGTGCAAAGTGAAGGTGAAGGA	450
Qy	1528	CCAGGCCCATGATGCCACTCATCAGGAGCTCTTAAGGCCCCAGGTAAAGTGCACAGTGAC	1587
Db	451	CCAGGCCCATGATGCCACTCATCAGGAGCTCTTAAGGCCCCAGGTAAAGTGCACAGTGAC	510
Qy	1588	AGATAAGGGTGTGAAAGGTCACTCTCGAGTGGGCAGGTGGGGTAGGGAAAGGCAAGGT	1647
Db	511	AGATAAGGGTGTGAAAGGTCACTCTCGAGTGGGCAGGTGGGGTAGGGAAAGGCAAGGC	570
Qy	1648	CATGTTCTGGAGAGGGGTGTGACTACATTAGGGTGTATGAGCCTAGCTTGGAGAGTGA	1707
Db	571	CATGTTCTGGAGAGGGGTGTGACTACATTAGGGTGTATGAGCCTAGCTTGGAGAGTGA	630
Qy	1708	TGCGCGGTCCACTGAGACCCCTGGTTATCCAGAAGCCTGTGTGGGCTTGGGAGAGCTTGG	1767
Db	631	TGCGCGGTCCACTGAACCCCTGGTTATCCAGAAGCCTTTCAGAGGCTTTCAGAGAGCTTGG	690
Qy	1768	AGTGGGGAGAGGGGTGACTTTCTCGACCAAGGCTTTTCTACCAACCTTACCCCTGGGTAAAGG	1827
Db	691	AGTGGGGAGAGGGGTGACTTTCTCGACCAAGGCTTTCTACCAAGGCTTTACCCCTGGGTAAAGG	750
Qy	1828	GCTGTGAGCAGGAAGCAGCGGGCAAGACCTCTGGAGCAGCCCATATCTGCGCTTGGCCTGA	1887
Db	751	GCCTGAGCAGGAAGCAGCGGGCAAGAACCTCTTGGAGCAGCCCATATCTGCGCTTGGCCTGA	810
Qy	1888	CTCTGCCACTGGCAGCACAGTCAACAGCAGGTTCACCTCACAGAGAGGGCGCAAGGCCA	1947
Db	811	CTCTGCCACTGGCAGCACAGTCAACAGCAGGTTCACCTCACAGAGAGGGCGCAAGGCCA	870
Qy	1948	TCATCAGCTCCCTTTATAAGGGAAGGGTCAACGGCTCGGTGTGTCGCGAGAGTGTCTTGCCT	2007
Db	871	TCATCAGCTCCCTTTATAAGGGAAGGGTCAACGGCTCGGTGTGTCGCGAGAGTGTCTTGCCT	930
Qy	2008	GGTCTCTGTGCTGTGGGTGGGGGTGCGCAGGTGTGTCCAGAGAGGCCCATAGTTGGTAG	2067
Db	931	GGTCTCTGTGCTGTGGGTGGGGGTGCGCAGGTGTGTCCAGAGAGGCCCATATTTGGTAG	990
Qy	2068	TGAGGCAGCCATGGGGCTAGAGCACTGTGTGCCCTTGGCCATCATATAGTGGCCATCTTCTCT	2127
Db	991	TGAGGCAGGTATGGGGCTAGAGCACTGTGTGCCCTTGGCCGTGATATAGTGGCCATCTTCTCT	1050
Qy	2128	GCTCTGTGTGGACCTCATGCAACGGCAACAAAGCTGGGCTGACAGCTTACCCGCCAGGTCT	2187
Db	1051	GCTCTGTGTGGACCTCATGCAACGGCGCCAAAGCTGGGCTGACAGCTTACCCAGGCCCC	1110
Qy	2188	CTTGCCACTGCCCGGCTGGGCAACCTTGTGTGATGTGACTTTCAGAAACACACCATACT	2247
Db	1111	CTTGCCACTGCCCGGCTGGGCAACC-TGCTGATGTGGACTTTCAGAAACACACCATACT	1169
Qy	2248	GCTTCCACAGGTGAGGAGAGGTCTCTGGAGGGCGCAGAGGTCTCTGAGATGCCCCAC	2307
Db	1170	GCTTCCACAGGTGAGGAGAGGTCTCTGGAGGGCGCAGAGGTCTCTGAGGTCTCCCTTAC	1229

2308 CACCAGCAACATGGGTGGTGTAAACCAACAGGCTGGATCAGAAAGCCAGGCTCAGAAG 2367
2309 CAGAGCAACATGGATGGTGGTGAACCAACAGGCTGACAGAAAGCCAGGCTGAGAAG 1289
2310 CAGAGCAACATGGATGGTGGTGAACCAACAGGCTGACAGAAAGCCAGGCTGAGAAG 1289
2368 GGGAGACAGGTTTGGGGGACGTTTCTCTGGGGAAGGACATTTATATATGATGGAATGAAGACTG 2427
2369 GGGAGACAGGTTTGGGGGACGTTTCTCTGGGGAAGGACATTTATATATGATGGAATGAAGACTG 2427
2390 GGGAGACAGGTTTGGGGGACGTTTCTCTGGGGAAGGACATTTATATATGATGGAATGAAGACTG 1348
2428 GATTTTCCAAAGCCAAAGAGTAGGCAAGAGGCTGGAGTGGAGTGGACTTGGCA 2487
2429 GATTTTCCAAAGCCAAAGAGTAGGCAAGAGGCTGGAGTGGAGTGGACTTGGCA 2487
2430 GATTTTCCAAAGCCAAAGAGTAGGCAAGAGGCTGGAGTGGAGTGGACTTGGCA 1408
2488 GTGGGCAATCAAGCCCAATTTGGGCAACATATGTTATGGAGTCAAAAGTCCCTTCTGCTGAC 2547
2489 GTGGGCAATCAAGCCCAATTTGGGCAACATATGTTATGGAGTCAAAAGTCCCTTCTGCTGAC 1468
2548 ACCAGAGAAAGGCTTGGGAATGGAAGATGAGTTAGTCTCTGAGTGGCGTTTAAATCAC 2607
2549 ACCAGAGAAAGGCTTGGGAATGGAAGATGAGTTAGTCTCTGAGTGGCGTTTAAATCAC 1528
2608 GAAATCGAGGATGAAGGGGCTGACGTGACCCGCTTCAACCTTTTGCACGTGGGTCTCTC 2667
2609 GAAATCGAGGATGAAGGGGCTGACGTGACCCGCTTCAACCTTTTGCACGTGGGTCTCTC 1588
2668 GGGGCTCACTG-CTCAACGGGATGACCATCATCTGGGAATGGGATGCTTAACTGGGGCT 2726
2669 GGGGCTCACTGCTCAACGGGATGACCATCATCTGGGAATGGGATGCTTAACTGGGGCT 1648
2727 CTGGCAATTTTGGTGAATCTTCAAGGTGATCTCTGGGATGAGTCTGAGTGGCGTTTAAATCAC 2786
2728 CTGGCAATTTTGGTGAATCTTCAAGGTGATCTCTGGGATGAGTCTGAGTGGCGTTTAAATCAC 1708
2787 CTTCATCAGAAAGGTGACCCGCTGACCCGCTTCAACCTTTTGCACGTGGGTCTCTC 2846
2788 CTTCATCAGAAAGGTGACCCGCTGACCCGCTTCAACCTTTTGCACGTGGGTCTCTC 1768
2847 CTTCACCTGCTCACTCTCTGTTAGCCCGGGGTGCTCAAGGTTCAAATAGGACTAGGA 2906
2848 CTTCACCTGCTCACTCTCTGTTAGCCCGGGGTGCTCAAGGTTCAAATAGGACTAGGA 1828
2907 CTGTAGTCTGGGTGATCTCTGGTGAACAGAGGCTTCAACCTCTCTGAGTGGCG 2966
2908 CTGTAGTCTGGGTGATCTCTGGTGAACAGAGGCTTCAACCTCTCTGAGTGGCG 1888
2967 GCGGCTTCCGGGACGTTTCAAGGTGATCTCTGGTGAACAGAGGCTTCAACCTCTCTGAGTGGCG 3026
2968 GCGGCTTCCGGGACGTTTCAAGGTGATCTCTGGTGAACAGAGGCTTCAACCTCTCTGAGTGGCG 1948
3027 TGGGCTGGGCGGCTGCGGAGCGATGGTGAACCGGCGGAGGACACCGGCGGAGCGGCG 3086
3028 TGGGCTGGGCGGCTGCGGAGCGGCTGGTGAACCGGCGGAGGACACCGGCGGAGCGGCG 2008
3087 GCCTGGGCGGCTGCGGAGCGGCTGGTGAACCGGCGGAGGACACCGGCGGAGCGGCG 3146
3088 GCCTGGGCGGCTGCGGAGCGGCTGGTGAACCGGCGGAGGACACCGGCGGAGCGGCG 2067
3147 GGGGACAGAGCGGCTTCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGG 3206
3148 GGGGACAGAGCGGCTTCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGG 2126
3207 CGGACAGAGCGGCTGCGGAGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGG 3266
3208 CGGACAGAGCGGCTGCGGAGCGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGG 2173
3267 ACAGTGGGCGGAGGAGGCTGCGGAGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGG 3318
3268 ACAGTGGGCGGAGGAGGCTGCGGAGGCTGGGCGGCTGGGCGGCTGGGCGGCTGGGCGG 2233
3319 -----GAGGGGGGGGGGCTACTGCGGAGGCTGGGCGGCTGGGCGGCTGGGCGG 3370
2234 TTGTGACGAGTGGGGGGGCGGCTGCGGAGGCTGGGCGGCTGGGCGGCTGGGCGG 2293
3371 -----ATGCGTGAAGTGGGGGGGCGGCTGGGCGGCTGGGCGGCTGGGCGG 3413

2294 CGCATTTTCCAGCTCGAATTCGGTGTCTGAAAGTGGGGCGGGACCGCACCTGTGCTGTAA 2353
3414 GCTCAGTGTGGCGGGACGG-----GCGGGAATCTCTTCTTGTAGTGAAGAGGTGGTCAAGGT 3468
2354 GCTCAGTGTGGGTGGCGGGGCGGCGGGGCTTCTCTCTGAGTGCAGAGGCGGTCAAGGT 2413
3459 GGGCAGAGACAGAGTGGGGCGCAAAACCGCCAGGAGGGGAGCAATGTGTGGGTGAGCAA 3528
2414 GGGCAGAGACAGAGTGGGG-CAAAGCTTGGCCAGCAAGGAGCAAGTGTGATGCACAA 2472
3529 AGAGTGGGCGCTGTCTCCAGCTTGGACCGGGCTAGGGAATGCGGGGAGACCTTGTGGAGCGC 3588
2473 AGAGTGGGCGCTGTCTCCAGCTTGGACCGGGCTAGGGAATGCGGGGAGACCTTGTGGAGCGC -A 2531
3589 CAGGTTTGGAGTGGGTGGCGGAGGAGTGGGCGCAAGGCTTCTATGCAAGCCCAAGTGTGTC 3648
2532 TAGGGTTGGAGTGGGTGGATGGTGGGGCTTAATGCTCTTATGTCGCAAGCGCACGTGCC 2591
3649 CGTCCGCGCCCGCAGGGGTGATCTCTGCGCTATGCGGCGCGGCTGCGGAGCAGAGGCG 3708
2592 CGTCCCAACCGGAGGGGTGCTCTGGCGGCTATGCGGCGGCTGCGGAGCAGAGGCG 2651
3709 CTTCTCCGTGTCACCTTTCGCAACTTGGGCGCTGGGCAAGAGTCTGCTGGAGCAGTGGGT 3768
2652 CTTCTCCGTGTCACCTTTCGCAACTTGGGCGCTGGGCAAGAGTCTGCTGGAGCAGTGGGT 2711
3769 GACCGAGAGGCGCGCTCTCTTGTGGCG-CTTTCGCGCAACGAGCGGCTGGGTGATGGCG 3827
2712 GACCGAGAGGCGCGCTCTCTTGTGGCGCTTTCGCAACTTGGGCGGCTGGGTGATGGCG 2771
3828 AGAGGGCACAAGCGGGAATGCGGAGGCGGGGACCGGAGGCAACCTCTTACCCCGC 3887
2772 AGAGGGCACAAGCGGGAATGCGGAGGCGGGGACCGGAGGCAACCTCTTACCCCGC 2831
3888 ATCTCCCGCACAAGCGGCGCTTTCGCGCGCTTTCGCAACTTGGGCAAGAGCAGTGGAGC 3947
2832 ATCT-CCCGCACAAGCGGCGCTTTCGCGCGCTTTCGCAACTTGGGCAAGAGCAGTGGAGC 2890
3948 AACGTGATCGCTCTCTCTGCGGCGCGCTTTCGAGAGTACGAGCAGTCTGCTCTCTC 4007
2891 AACGTGATCGCTCTCTCTGCGGCGCGCTTTCGAGAGTACGAGCAGTCTGCTCTCTC 2950
4008 AGGCTGCTGGAGCTAGCTCAGAGGAGCTGAAAGAGGAGTGGGCTTCTTGGCGGAGGTG 4067
2951 AGGCTGCTGGAGCTAGCTCAGAGGAGCTGAAAGAGGAGTGGGCTTCTTGGCGGAGGTG 3010
4068 CGGAGCGAGAGACCGAGGAGTCTCTGCGGCGGCTTCTGAGAGTGGCGGCTGGAGC 4127
3011 CGGAGCGAGAGACCGAGGAGTCTCTGCGGCGGCTTCTGAGAGTGGCGGCTGGAGC 3070
4128 TGGGCGCTCCGAGAGGCGAGATTTGATAGTGGTGGGAAAGGACATTCAGAGAGC 4187
3071 TGGGCGCTCCGAGAGGAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 3130
4188 CCCACTGTAAAGAGGCGCTGGAGGAGGAGGAGTCTCAGACATGCTGCTGGGAGAGGT 4247
3131 CCCACTGTAAAGAGGCGCTGGAGGAGGAGGAGTCTCAGACATGCTGCTGGGAGAGGT 3190
4248 GTGCGCGGCTCAGAGGCGGAGCAGAGAGGCGCAAGGAGTCTGTAAC-CCGCTCCAGGTTGG 4306
3191 GTGCGCGGCTCAGAGGCGGAGCAGAGAGGCGCAAGGAGTCTGTAACCTCTTATCCAGTCA 3250
4307 AGATTTTCGATTTTGTGTTTCTCTCTGGGCAAGGAGAGGAGTGGAGTGGAGTGGAG 4366
3251 AGATTTTCGATTTTGTGTTTCTCTCTGGGCAAGGAGAGGAGTGGAGTGGAGTGGAG 3310
4367 GAGGAGCTTGGTGGTGGTAAAGGACAGGAGGCGGCTGGGCTACCTGGAGATGCG 4426
3311 GAGGAGCTTGGTGGTGGTAAAGGACAGGAGGCGGCTGGGCTACCTGGAGATGCG 3370
4427 TGGGCGCTCAGATTTGTCTCAGGTGAACGAGAGCAGGAGGAGTGGAGTGGAGTGGAG 4486

3371	Db	TGGGGCCTGAGACTTGTCCAGGTGAACGCGAGACACAGAGAGGAAATTGAGACCCCGTTCTG	3430	QY	5554	GCTGSCCTGTCCATCCAGAGCCCGCTCTAGTGGGGAGACAAACAGGACCTGCGCAGAAT	5613
4487	QY	TCCTGGTGTAGGTCTGAATGCTGTCCCGCTCCTCTGACACATCCCGAGCGCTGGCTGGCAA	4546	Db	4509	GCTGSCCTGTCCATCCAGAGCCCGCTCTAGTGGGGAGACAAACAGGACCTGCGCAGAAT	4568
3431	Db	TCCTGGTGTAGGTCTGAATGCTGTCCCGCTCCTCTGCAATCCCGAGCGCTGGCTGGCAA	3490	QY	5614	GTTTGAGAGACCCAGCGCCTGACGGAGAGGGGGGAGTGTGGTGCTCTGAGAGGTGTGA	5673
4547	QY	GCTCTACGTTTCCAAAAGGCTTCTTGACCCAGCTGGAATGAGTGTCTAACTGAGCAG	4606	Db	4569	GTTTGAGAGACCCAAACGCTGACGGAGAGGGGGGAGTGTGGTGCTCTGAGAGGTGTGA	4628
3491	Db	GCTCTACGTTTCCAAAAGGCTTCTTGACCCAGCTGGAATGAGTGTCTAACTGAGCAG	3550	QY	5674	CTGGCCCTGTCTGTGGGGTCCGAGAGGGTACTGTGGAGCTTCTCGGCGCAGAGACTAGTT	5733
4607	QY	GATGACCTGGGACCCAGCCAGCCACCGGAGACCTGACTGAGCGCTTCTCGGCAAGAA	4666	Db	4629	CTGGCCCTGTCTGTGGGGTCCGAGAGGGTACTGTGGAGCTTCTCGGCGCAGAGACTAGTT	4688
3551	Db	GATGACCTGGGACCCAGCCAGCCACCGGAGACCTGACTGAGCGCTTCTCGGCAAGAT	3610	QY	5734	GACAGAGTCCAGTGTGTGTCGCCAGGCAGTGTGTCTCCCGTGTGTTTGGTGCGCAGGGGTC	5793
4667	QY	GGAAGAGGTGAGAGTGGCTGCCACGCTGGGGGGCAAGGCTGGTGGTTGAACGTCCACAGG	4726	Db	4689	GACAGAGTCCAGTGTGTGTCGCCAGGCAGTGTGTCTCCCGTGTGTTTGGTGCGCAGGGGTC	4748
3611	Db	GGAAGAGGTGAGAGTGGCTGCCACGCTGGGGGGCAAGGCTGGTGGTTGAGGCTCCACAGG	3670	QY	5794	CCAGATCTGAGAGTCCAGTCCCACTCTCAACCTGATCTCTGTCGCCAGGGAACGACAC	5853
4727	QY	AGGAATGAGGGAGGCTGGGCAAAAGTTGGAACAGTGCATCAACCGCGCAGCCGATCT	4786	Db	4749	CCAGATCTGAGTCCAGTCCCACTCTCAACCTGATCTCTCTGTCGCCAGGGAACGACAC	4808
3671	Db	AGGAATGAGGGAGGCTGGGCAAAAGTTGGAACAGTGCATCAACCGCGCAGCCGATCT	3730	QY	5854	TCATCAACCAACCTGCTGAGAGGTGAGGCGCTCTGGGAGAACCCCTTCCGCT	5913
4787	QY	GGGCTGACAGGTGCAAAATTGAGAGTCAATTTGGGGGCTACCCCGTCTATCCCTGAGTA	4846	Db	4809	TGATCAACCAACCTGCTGAGAGGTGAGGCGCTCTGGGAGAACCCCTTCCGCT	4868
3731	Db	GGGCTGACAGGTGCAAAATTGAGAGTCAATTTGGGGGCTACCCCGTCTATCCCTGAGTA	3788	QY	5914	TCCACCCCGAAACACTTCTCGGATGCCAGGGCCACTTTGTGAAGCCGAGGCGCTTCTGTC	5973
4847	QY	TCCTCTCGGCGCTGCTCAGGCCAAAGGGAGCCCTGAGAGCAGCTTCAATGATGAGAACCT	4906	Db	4869	TCCACCCCGAAACACTTCTCGGATGCCAGGGCCACTTTGTGAAGCCGAGGCGCTTCTGTC	4928
3789	Db	TGCTCTCGGCGCTGCTCAGGCCAAAGGGAGCCCTGAGAGCAGCTTCAATGATGAGAACCT	3848	QY	5974	CTTCTCAGCAGGTGCTGTGGGAGACCCGCGCTCCCTGTCCCTTCCGTGGAGTCTTGCA	6033
4907	QY	GCGCATAGTGGTGGTAACTGTTCTTGTGCGGGATGTTGACCACTCGACCAACGCTGGC	4966	Db	4929	CTTCTCAGCAGGTGCTGTGGGAGACCCGCGCTCCCTGTCCCTTCCGTGGAGTCTTGCA	4988
3849	Db	GCGCATAGTGGTGGTAACTGTTCTTGTGCGGGATGTTGACCACTCGACCAACGCTGGC	3908	QY	6034	GGGGTATCACCCAGGAGCCAGCTCAGTGACGCCCTCCCTCCCGCAGCGCCGCTGC	6093
4967	QY	CTGGGCGCTCTGCTCATGATCCTACACTGATGTGACGTGAGCCAGCTGGGCGCC	5026	Db	4989	GGGGTATCACCCAGGAGCCAGCTCAGTGACGCCCTCCCTCCCGCAGCGCCGCTGC	5048
3909	Db	CTGGGCGCTCTGCTCATGATCCTACACTGATGTGACGTGAGCCAGCTGGGCGCC	3968	QY	6094	ATGCGCTCGGGAGAGCCCTGGCGCCGATGAGAGCTCTTCTCTCTTCTTCACTCCCTGCTGCA	6153
5027	QY	AAGGACAGGACTGAGGGAGGAGGTACAGCTGGGGGCGCTTGGGCTTAGCTGGGACACC	5086	Db	5049	ATGCGCTCGGGAGAGCCCTGGCGCCGATGAGAGCTCTTCTCTCTTCTTCACTCCCTGCTGCA	5108
3969	Db	AGTGAGGGGCGAGGGAGGAGGTAAGGGCGGGGCGCCATGAACTTTGCTGGGACACC	4028	QY	6154	GCATTTAGCTTCCGFTGGCGCGGAGAGCCCGGCGCAGCCCTCTCGTGTGTCAG	6213
5087	QY	CGGGGCTTCCAGCAGCGGTGGCGAGGCTCTGTAAGCCTAACTTCTCTCAACACAGGA	5146	Db	5109	GCATTTAGCTTCCGFTGGCGCGGAGAGCCCGGCGCAGCCCTCTCGTGTGTCAG	5168
4029	Db	CGGGGCTTCCAGCAGCGGTGGCGAGGATCTCTGAAGCCTGACCTCTCTCAACATAGGA	4088	QY	6214	CTTCTGTGTGACCCCATCCCGCTACGAGCTTTGTGCTGCCCCCGCTAGATGGGGTACC	6273
5147	QY	GGAAAGAGAGTGT-----CCCTGGGTGCTGACCCATTTGGGGAGCGCATGT	5193	Db	5169	TTTCTGTGTGAGCCCATCCCGCTATGAGCTTTGTGCTGCCCCCGCTAGATGGGGTACC	5228
4089	Db	GGAAAGAGAGTGT-----CCCTGGGTGCTGACCCATTTGGGGAGCGCATGT	4148	QY	6274	TAGTCTCCAGCGCTGCTCCCTAGCCAGAGGCTCTAATGTACAAATAAGCAATGTGGTATGT	6333
5194	QY	CTGTCCAGTCCGTTCCAAACAGAGATCGACGAGCTGATAGGGAGGTGCGGCGACACGA	5253	Db	5229	TAGTCTCCAGCGCTGCTCCCTAGCCAGAGGCTCTAATGTACAAATAAGCAATGTGGTATGT	5288
4149	Db	CTGTCCAGGCGGTGTCACACAGAGATCGACGAGCTGATAGGGAGGTGCGGCGACACGA	4208	QY	6334	CCAACTTGGGTGCTGTCACGCGCTGCTGTTGGGATCATCTCTCTAGGGGCAACCCACCC	6393
5254	QY	GATGGGTGACCAAGGCTCAATGCCCTTACACCACTCCCGTGAATCACAGGTTGACGCGCTT	5313	Db	5289	CCAACTTGGGTGCTGTCACGCGCTGCTGTTGGGATCATCTCTCTAGGGGCAACCCACCC	5348
4209	Db	GATGGGTGACCAAGGCTCAATGCCCTTACACCACTCCCGTGAATCACAGGTTGACGCGCTT	4268	QY	6394	CTGTGCTCATTTCTGCTTACCCCAACCGCTGCGCGCATTTTGAGAC--GGGTACGCTTGGG	6451
5314	QY	TGGGGACATCATCCCGCTGAGTGTGACCCATATGACATCCCGTGAATCGAGGTACAGGG	5373	Db	5349	CTGTGCTCATTTCTGCTTACCCCAACCGCTGCGCGCATTTTGAGAC--GGGTACGCTTGGG	5408
4269	Db	TGGGGACATCATCCCGCTGAGTGTGACCCATATGACATCCCGTGAATCGAGGTACAGGG	4328	QY	6452	CTGAGCAGATGTCAAGTTACCTTGGCCATAATCCCATGTCCCCCACTGACCCCACTCTGGA	6511
5374	QY	CTTCCGATCTTAAAGTTAGGCTTGGCGGCTTCTCAGCCAGCTCAGACACGACCTG	5433	Db	5409	CTGAGCAGATGTCAAGTTACCTTGGCCATAATCCCATGTCCCCCACTGACCCCACTCTGGA	5468
4329	Db	CTTCCGATCTTAAAGTTAGGCTTGGCGGCTTCTCAGCCAGCTCAGACACGACCTG		QY	6512	CTGCCCAATTGGTGACAAAGACTACATTTGCTTGGCATGTGGGGAAAGGGGCGCAGAAATGG	6571
5434	QY	GTGATAGCCCCCAGCATGGGTACTGCCAGGTGGGGCCCACTCTAGGAACCCCTGGCCACCTAG	5493	Db	5469	CTGCCCAATTGGTGACAAAGACTACATTTGCTTGGCATGTGGGGAAAGGGGCGCAGAAATGG	5528
4389	Db	GTGATAGCCCCCAGCATGGGTACTGCCAGGTGGGGCCCACTCTAGGAACCCCTGGCCACCTAG	4448	QY	6572	GCTGACTAGAGGTGTCAAGCTCAGCCCTGGATGTGTGGAGAGGGGCGAGGACTCAGCGCTGGAG	6631
5494	QY	TCCTCAATGCCACCACTGACTGTGCCCACTTGGGGTGGGGGGTCCAGAGTATAGGCAGG	5553	Db	5529	GCTGACTAGAGGTGTCAAGCTCAGCCCTGGATGTGTGGAGAGGGGCGAGGACTCAGCGCTGGAG	5588
4449	Db	TCCTCAATGCCACCACTGACTGTGCCCACTTGGGGTGGGGGGTCCAGAGTATAGGCAGG	4508				

Qy	6632	GCCCATATTTTCAGGCGCTAACTCAGGCCACCCACATCAGGGACGAGTCTCTGCCAGCAC	6691
Db	5589	GCCCATATTTTCAGGCGCTAACTCAGGCCACCCACATCAGGGACGAGTCTCTGCCAGCAC	5648
Qy	6692	CATCACAAACAGTCACCTCCCTTCATATATGACACCCCAAAATCGAGACAAAATCAT---	6748
Db	5649	CATCACAAACAGTCACCTCCCTTCATATATGACACCCCAAAACGGAAGACAAATCATGGCG	5708
Qy	6749	TCAGGGAGCTATATGCCAGGG-----CTACCTTCCCGAGGCTCAGTCGGCAGGTGCCAGAAC	6804
Db	5709	TCAGGGAGCTATATGCCAGGGCTACCTACCTCCAGGGCTCAGTCGGCAGGTGCCAGAAC	5768
Qy	6805	ATTCCCTGGGAAGGCCCCAGGAAACCCAGGACCGACCCACCGCCCTTCAGCCT	6857
Db	5769	GTTCCCTGGGAAGGCCCCATGGAAGCCAGGACTGAGCCACCACTTCAGCCT	5821

RESULT 7	
ADB25775	
ID	ADB25775 standard; DNA; 4500 BP.
XX	
AC	ADB25775;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Human CYP2D6-related DNA sequence.
XX	
KW	human; mutant CYP2D6 gene; drug analysis; drug testing; gene; ds.

QY	2258	GGTGAAGGAGAGGTCTCTGAGGCGCGCAGAGGTCTCTGAGGATGCCGCCACACACGAAA	2317
DB	180	GGTGAAGGAGAGGTCTCTGAGGCGCGCAGAGGTCTGAGGCTCCCTACCAAGACAAA	239
QY	2318	CATGGGTGGTGGGTTAAACACACAGGCTGGATCAGAAGCCAGGCTGAGNAGGGGAAGCAGG	2377
DB	240	CATGATGGTGGGTGAAACACACAGGCTGGACCAAGAGCCAGGCTGAGNAGGGGAAGCAGG	299
QY	2378	TTTGGGGGACGTTCTCTGGGAAGGACATTTATATCATGGCATGAAGACTTGGATTTTCCAA	2437
DB	300	TTTGGGGGACG-TCCTGGAGAGGGCAATTTATACATGGCATGAAGACTTGGATTTTCCAA	358
QY	2438	AGGCCAAGGAAGATAGGGCAAGGGCTGGAGGTGGAGCTGGACTTGGCAGTGGGCATGC	2497
DB	359	AGGCCAAGGAAGATAGGGCAAGGGCTGGAGGTGGAGCTGGACTTGGCAGTGGGCATGC	418
QY	2498	AAGCCCCATTTGGGCAAATATGTTATGAGGATPACAAAGTCCCTTCTGCTGCACACAGAAAGGA	2557
DB	419	AAGCCCCATTTGGGCAAATATGTTATGAGGATPACAAAGTCCCTTCTGCTGCACACAGAAAGGA	478
QY	2558	AAGGCCTTTGGGAATGGAAGATGATTAGTCCCTCAGTGGCGTTTAAATCAGAAATCGAGG	2617
DB	479	AAGGCCTTTGGGAATGGAAGATGATTAGTCCCTCAGTGGCGTTTAAATCAGAAATCGAGG	538
QY	2618	ATGAAGGGGGTGCAGTGACCCGGGTTCAAACTTTTGCACCTGTGGGTCTCTCGGSCCTCACT	2677
DB	539	ATGAAGGGGGTGCAGTGACCCGGTCAAACTTTTGCACCTGTGGGTCTCTCGGSCCTCACT	598
QY	2678	G-CTCACCGGATGGACCATCATCTCGGGAATGGGATGCTTAACCTGGGGGCTCTCTCGGCAATT	2736
DB	599	GCCTCAACCGGATGGACCATCATCTCGGGAATGGGATGCTTAACCTGGGGGCTCTCTCGGCAATT	658
QY	2737	TTGGTGACTCTTGCAGAGGTATACCTCGGGTGACGCATCCAAACTGATGTTCTCTCCATCACA	2796
DB	659	TTGGTGACTCTTGCAGAGGTATACCTCGGGTGACGCATCCAAACTGATGTTCTCTCCATCACA	718
QY	2797	GAAGGTGTGACCCGCCACCCCTGCCCCACCATCAGGAGGCTGGGTCTCTCTTCCACCTG	2856
DB	719	GAAGGTGTGACCCGCCACCCCTGCCCCACCATCAGGAGGCTGGGTCTCTCTTCCACCTG	778
QY	2857	CTCATCTCTGGTAGCCCCGGGGTCTGTCCAAGGTTCAATATAGGACTTAGGACCTGTAGTCT	2916
DB	779	CTCATCTCTGGTAGCCCCGGGGTCTGTCCAAGGTTCAATATAGGACTTAGGACCTGTAGTCT	838
QY	2917	GGGTGTACTCTGGCTTGACAGAGGCGCTGACCCCTCCTCTGCAGTTGGCGCGCCGCTTC	2976
DB	839	GGGTGTACTCTGGCTTGACAGAGGCGCTGACCCCTCCTCTGCAGTTGGCGCGCCGCTTC	898
QY	2977	GGGACCTGTTACGCTTGACGCTGGCTGGACGCCCGTGGTCTGTCTCAATGGGCTGGCG	3036
DB	899	GGGACCTGTTACGCTTGACGCTGGCTGGACGCCCGTGGTCTGTCTCAATGGGCTGGCG	958
QY	3037	GCCGTGCGAGGCGATGTTGACCCCGCGCGAGGACACCGGCCGACCCCGCTGCGCCCC	3096
DB	959	GCCGTGCGAGGCGCTGTGTGACCCAGCGGAGGACACCGGCCGACCCCGCTGCGCCCC	1018
QY	3097	ATCTACAGAGTCTTGGGCTTCGGGGCGCGTTTCCAAAGGACAGCGGCGGTGGGGACAGAG	3156
DB	1019	ATCACCCAGATCTCTGGGTTTTTGGGCGCGGTTCCCAAGGCAAGCAGCGGT- GGGGACAGAG	1077
QY	3157	ACGCGTTTCCGTGGGCGCCGGGTGGACAGTACCCTAGCCCCAAGCAGCGCCGACAGGGC	3216
DB	1078	ACAG-ATTTCCGTGGAGACCCGGGTGGGTGATGACCGTTAGTTCGAGCTGGCGAGAGGGC	1136
QY	3217	GTGGGGTCTCTGAGCTGAAACAGAGATAAAGGCCAGCGAGTGGGCTGAGGACAGTGGGCC	3276
DB	1137	CGGGGTCTGGACATGAAC-----AGGCCAGCGAGTGG-----GGAACAGCGGGC	1183
QY	3277	AGGAAACCACTTGACGGGGGAGGTCGAGTCTGTGGGCTGG-----GAG	3321
DB	1184	AAGAAACCACTTGACCTTAGGGAGGTGTGACATGTGGGGAAGAGGGCGGGGCTGTGTGACGAG	1243

QY 3322 GGGCGGGGCTACTCCAGACCCCGCAGAAAGCCCGGTGGCGAGGCTG----- 3370
DB 1244 TGGGCGGGGCCACTGCCAGACCTGGCAGAGGCCCAATGGGTGAGCGTGGCGCATTTCCC 1303
QY 3371 -----ATGCGTCGAAGTGGCGGTGGCGGGACCGGCCCTATGCTGCGGGCTCAGTGTG 3423
DB 1304 AGCTGGNATCCGGTGTGCAAGTGGGGCGGGGACCGCACCTGTGCTGTAAAGCTCAGTGTG 1363
QY 3424 G-----GCGGGA CCGGGCGGATTTCTTCTCAGTGGAAAGTGTGTCAGGGTGGGCAGAGAC 3478
DB 1364 GGTGGCGCGGGGCCCGCGGGTCTTCTCCTCAGTGCAAAGCGGTTCAGGGTGGGCAGAGAC 1423
QY 3479 GAGGTGGGCAAAACCCCGCCAGCAGCGGGAGCAATGTGGTGGACAAAGAGTGGGCC 3538
DB 1424 GAGGTGGG--CAAGGCCCTGCCAGCCCAAGGAGCAAGGTGGATGACAAAGAGTGGGCC 1482
QY 3539 CTGTGCCAGCTGGACCGGGCTAGGGAATCGCGGAGACCTTGTGGAGCGCCAGGGTTGGA 3598
DB 1483 CTGTGACCAGCTGGACAGACCGAGGACTCGGGAGACCAAGGGGAGC--ATAAGGTTGGA 1541
QY 3599 GTGGGTGGCGAGGGTGGGGCCAAAGGCTTTCATGGCAACGCCACGTCGTCCTCCGCCCC 3658
DB 1542 GTGGGTGGTGGTGGGGCTAATGCTTTCATGGCCACGCGCACGTCGCCCGTCCCAACC 1601
QY 3659 CAGGGGTGATCTGTGCGCTATGGCCCGCGTGGCGGCGAGCAGAGCGCTTCTCCGTG 3718
DB 1602 CAGGGGTGTTCCTGGCGCTATAGGCCCGCGTGGCGGAGCAGAGCGGCTTCTCCGTC 1661
QY 3719 TCACACTTGGCAACTTGGGCTGGGCAAGTGTCCGTGGAGCAGTGGGTGACCGAGGAG 3778
DB 1662 TCACACTTGGCAACTTGGGCTGGGCAAGTGTCCGTGGAGCAGTGGGTGACCGAGGAG 1721
QY 3779 GCGGCTGCTTGTGCGG--CTTCGCGCAACCAAGCCGGTGGGTGATGGGCAGAAAGGCGAC 3837
DB 1722 GCGGCTGCTTGTGCGGCTTCGCGCAACCACTCCGGTGGGTGATGGGCAGAAAGGCGAC 1781
QY 3838 AAAGCGGGAJCTGGGAAGCGGGGACGGAAGAGCAACCCCTTACCGGCATCTCCCCAC 3897
DB 1782 AAAGCGGGAJCTGGGAAGCGGGGACGGAAGAGGCGACCCCTTACCGGCATCT-CCCA 1840
QY 3898 CCCCAGAGCCCTTTCGCCCAACCGGCTCTTGGACAAGCCGTGGACCAAGTATCG 3957
DB 1841 CCCAGAGCCCTTTCGCCCAACCGGCTCTTGGACAAGCCGTGGACCAAGTATCG 1900
QY 3958 CCTCCCTCAGCTCGGGCGCGCTTCAGTACAGCAACCTCGCTTCTCAGGCTGCTGG 4017
DB 1901 CCTCCCTCAGCTCGGGCGCGCTTCAGTACAGCAACCTCGCTTCTCAGGCTGCTGG 1960
QY 4018 ACCTAGCTCAGGAGGACTGAAGAGAGTGGGCTTTTCGCGAGGTGCGGAGCGAGA 4077
DB 1961 ACCTAGCTCAGGAGGACTGAAGAGAGTGGGCTTTTCGCGAGGTGCGGAGCGAGA 2020
QY 4078 GACGAGGAGTCTTCGAGGGGAGCTCTGAGAGGTGCGGGCTGGAGTGGGGCTCC 4137
DB 2021 GACGAGGAGTCTTCGAGGGGAGCTTCGAGAGGTGCGGGGCTGGAGTGGGGCTCG 2080
QY 4138 GAAGGGCAGATTTGCATAGATGGGTTTGGGAAGGACATTTCCAGGAGACCCCACTGTAA 4197
DB 2081 GAAGAGCAGATTTGCATAGATGGGTTTGGGAAGGACATTTCCAGGAGACCCCACTGTAA 2140
QY 4198 GAAGGGCTCGGAGGAGGGGACATCTCAGACATGGTGTGGGAGAGGTGTGCCCGGT 4257
DB 2141 GAAGGGCTCGGAGGAGGGGACATCTCAGACATGGTGTGGGAGAGGTGTGCCCGGT 2200
QY 4258 CAGGGGCACAGGAGGCGCAAGGACTCTGTACC--CCGTCACGTTGGAGATTTTCGAT 4316
DB 2201 CAGGGGCACAGGAGGCGCAAGGACTCTGTATCTTATCTTATCCACGTCAGAGATTTTCGAT 2260
QY 4317 TTTAGGTTTCTCTCTGGGCAAGGAGAGGGTGGAGGCTGGCACTTTGGGAGGGACTTG 4376
DB 2261 TTTAGGTTTCTCTCTGGGCAAGGAGAGGGTGGAGGCTGGCACTTTGGGAGGGACTTG 2320
QY 4377 GTGAGGTCAGTGGTAAGGACAGGACGCGCCCTGGGTCTTACTGGAGATGGCTGGGGCTGA 4436

DB 2321 GTGAGGTCAGTGGTAAGGACAGGACGCGCCCTGGGTCTACTGGAGATGGCTGGGGCTGA 2380
QY 4437 GACTTCTCAGGTGAACCGCAGCAGCAGGAGGATTTGAGACCCCGTTCGTCTGTGTGTAG 4496
DB 2381 GACTTCTCAGGTGAACCGCAGCAGCAGGAGGATTTGAGACCCCGTTCGTCTGTGTGTAG 2440
QY 4497 GTGCTGAATGTGTCTCCCTCTCTCTGCAATCCAGCGCTGGCTGGCAAGGTCTTACGC 4556
DB 2441 GTGCTGAATGTGTCTCCCTCTCTCTGCAATCCAGCGCTGGCTGGCAAGGTCTTACGC 2500
QY 4557 TTTCCAAAAGGCTTTCTGTGACCCAGCTGGATGAGCTTAACCTGAGCACAGGATGACCTGG 4616
DB 2501 TTTCCAAAAGGCTTTCTGTGACCCAGCTGGATGAGCTTAACCTGAGCACAGGATGACCTGG 2560
QY 4617 GACCCAGCCAGCAGCCCGCAGACCTGACTGAGGCTTCTTGGCAAGAGAGGAAAGGTG 4676
DB 2561 GACCCAGCCAGCAGCCCGCAGACCTGACTGAGGCTTCTTGGCAGAGATGGAGAAAGGTG 2620
QY 4677 AGAGTGGCTGCCACGCTGGGGGGCAAGGCTGGGTGGTGGAACTCCAGGAGGAAAGG 4736
DB 2621 AGAGTGGCTGCCACGCTGGGGGGCAAGGCTGGGTGGTGGTGGTGGTGGTGGTGGTGG 2680
QY 4737 GGAGGCTGGGCAAAAGGTTGGACCACTGATGATGATGATGATGATGATGATGATGATGATG 4796
DB 2681 GGAGGCTGGGCAAAAGGTTGGACCACTGATGATGATGATGATGATGATGATGATGATGATG 2740
QY 4797 GTGCAGAA TTTGAGGTTCA TTTGGGGCTACCCGTTCTATCCCTCAGTATCTCTCTCGG 4856
DB 2741 GTGCAGAA TTTGAGGTTCA TTTGGGGCTACCCGTTCTCTGCTCC--GAGTATGCTCTCGG 2798
QY 4857 CCTGCTCAGGCCAAGGGGAGCCTTGAGAGCAGTTCATGATGAGAACTCGGCATAGTG 4916
DB 2799 CCTGCTCAGGCCAAGGGGAACTTGAGAGCAGTTCATGATGAGAACTCTGTCATAGTG 2858
QY 4917 GTGGGTAACTGTTCTTCTGCGGGATGGTGAACA CTTGAGCAACGCTGGGCTGGGGCTC 4976
DB 2859 GTGGGTGACTGTTCTTCTGCGGGATGGTGAACA CTTGAGCAACGCTGGGCTGGGGCTC 2918
QY 4977 CTGCTCATCATCTTACACCTGATGTGACGCTGAGCCAGCTGGGCGCCCAAGGAGGGA 5036
DB 2919 CTGCTCATCATCTTACACCTGATGTGACGCTGAGCCATCTGGGAAACAGTGTGAGGGG 2978
QY 5037 CTGAGGAGGAAAGGTTACAGCTGGGGGCCCTGGGCTTAGCTGGGACACCCGGGGCTTCC 5096
DB 2979 CCGAGGAGGAAAGGTTACAGCGGGGGCCCATGAACTTTGCTGGGACACCCGGGGCTCCA 3038
QY 5097 AGCAGAGGCTGGCAGGCTCTGTAAAGCTTAACTTTCTTCCAA CACAGGAGGAAAGAGAG 5156
DB 3039 AGCAGAGGCTTGACCAAGGATCTCTGTAAGCTTGAACCTCTTCCAA CATAAGGAGGCAAGAGG 3098
QY 5157 TGT-----CCCTGGGTGCTGACCCCATTTGGGGAGCGATCTGTCCAGTTC 5203
DB 3099 AGTGTGAGGGCGGACCCCTCTGGGTGCTGACCCATTTGGGGAGCGATGTGTGTCAGGC 3158
QY 5204 CGTGTCCAAACAGAGATCGACACCTGATAGGCGAGGTGCGGCGACAGAGATGGGTGAC 5263
DB 3159 CGTGTCCAAACAGAGATCGACACCTGATAGGCGAGGTGCGGCGACAGAGATGGGTGAC 3218
QY 5264 CAGGCTCA CATGCCCTTACCACTGCGCTGATTTACAGAGGTGAGCGCTTTTGGGGACATC 5323
DB 3219 CAGGCTCA CATGCCCTTACCACTGCGCTGATTTACAGAGGTGAGCGCTTTTGGGGACATC 3278
QY 5324 ATCCCTCAGTGTGACCCCATATGATCCCTGATCATCGAAGTACAGGGCTTCCGCATC 5383
DB 3279 GTCCCCCTGGGTGTGACCCCATATGATCCCTGATCATCGAAGTACAGGGCTTCCGCATC 3338
QY 5384 CTTAAAGTGAAGGCTGGCGCCCTTCCACCCAGCTCAGCACCCAGCACTCTGGTGTATAGCCC 5443
DB 3339 CTTAAAGTGAAGGCTGGCGCCCTTCCACCCAGCTCAGCACCCAGCACTCTGGTGTATAGCCC 3398
QY 5444 CAGCATGGCTACTGCGAGGTGGGCCCATCTTATAGGAACCCCTGAGCACTTATGCTCTAATGC 5503


```
QY 4398 GGCAGGCGCTGGGTCTACCTGGAGATGGCTGGGGCGCTGAGACTTGTCCAGGTGAACGCAG 4457
DB 181 GGCAGGCGCTGGGTCTACCTGGAGATGGCTGGGGCGCTGAGACTTGTCCAGGTGAACGCAG 240
QY 4458 AGCACAGGAGGGATTGAGACCCCGTTCTGTCTGTGTAGTGTCTGTAATGCTGTCCCGCTC 4517
DB 241 AGCACAGGAGGGATTGAGACCCCGTTCTGTCTGTGTAGTGTCTGTAATGCTGTCCCGCTC 300
QY 4518 CTCCTGCAATCCAGCGCTGGCTGGCAAGGTCCTACGTTTCCAAAGGCTTTCTCTGACC 4577
DB 301 CTCCTGCAATCCAGCGCTGGCTGGCAAGGTCCTACGTTTCCAAAGGCTTTCTCTGACC 360
QY 4578 CAGCTGATGAGCTGTCTAACTGAGCACAGATGACCTGGGACCCAGCCAGCCACCCCGA 4637
DB 361 CAGCTGATGAGCTGTCTAACTGAGCACAGATGACCTGGGACCCAGCCAGCCACCCCGA 420
QY 4638 GACCTGACTCAGGCTTCTCTGGCAAGAGGAGAGTGTGGCTGCTGCACCGGTGGG 4697
DB 421 GACCTGACTCAGGCTTCTCTGGCAAGAGTGTGGAGAGTGTGGCTGCTGCACCGGTGGG 480
QY 4698 GGCAAGGGTGGTGGGTGAA CGTCCAGGAGGAATGAGGGGAGGCTGGGCAAAAGGTTGG 4757
DB 481 GGCAAGGGTGGTGGGTGAGCGTCCAGGAGGAATGAGGGGAGGCTGGGCAAAAGGTTGG 540
QY 4758 ACCAGTGATCATCCGGCGAGCGCATCTGGGCTGACAGGTGAGAAATTGGAGGTCAATT 4817
DB 541 ACCAGTGATCATCCGGCGAGCGCATCTGGGCTGACAGGTGAGAAATTGGAGGTCAATT 600
QY 4818 GGGGGCTACCCCGTCTATCCCTGAGTATCTCTCGGCCCTCTCAGGCCAAGGGGAGC 4877
DB 601 GGGGGCTACCCCGTCTGTCTCC -- GAGTATGCTCTCGGCCCTCTCAGGCCAAGGGGAA 658
QY 4878 CCTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGTGGGTAACTGTTCTTGCC 4937
DB 659 CCTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGTGGGTAACTGTTCTTGCC 718
QY 4938 GGGATGTGACACCTGTCACACGCTGGCTGGGCTCTGTCTATGATCTTACACCTG 4997
DB 719 GGGATGTGACACCTGTCACACGCTGGCTGGGCTCTGTCTATGATCTTACACCTG 778
QY 4998 GATGTGAGCGTGTAGCCAGCTGGGCGCCAAAGCAGGAGTGTGGGAGGAGGATGACGC 5057
DB 779 GATGTGAGCGTGTAGCCAGCTGGGCGCCAAAGCAGGAGTGTGGGAGGAGGATGACGC 838
QY 5058 TGGGGGCGCTGGCTGTAGTGGGACACCCGGGCTTCCAGACAGCGCTGGGCGAGGCTC 5117
DB 839 CGGGGGCCCATGAATTTGCTGGGACACCCGGGCTTCCAGACAGCGCTTGAACAGGATC 898
QY 5118 CTGTAAAGCCTAACTTCTCCAAACAGGAGGAAGGAGTGT-----CCCT 5164
DB 899 CTGTAAAGCCTGACCTCTCCAAACATAGGAGGCAAGAGGAGTGTGAGGGCGGACCCCT 958
QY 5165 GGGTGTGACCCATTGTGGGAGCGCATGTCTGTCCAGTCCGTGTCNAAGGAGATGAC 5224
DB 959 GGGTGTGACCCATTGTGGGAGCGTGTCTGTCCAGGCGCTTCCAAAGGAGATGAC 1018
QY 5225 GAGCTGATAGGGCAGGTGGCGGACACAGAGATGGGTGACACAGCTCACATGCCCTACACC 5284
DB 1019 RAGCTGATAGGGCAGGTGGYGCACNAGATGGGTGACCGGCTRCATGCCCTCAVC 1078
QY 5285 ACTGCCGTGATTCACGAGGTGACGCGTTTGGGACATCATATCCCTGATGTGACCCAT 5344
DB 1079 ACTGCCGTGATTCAYGAGGTGACGCGTTTGGGACATCGTCCCTGGGTGTGACCCAT 1138
QY 5345 ATGACATCCCGTGACATCGAAGTACAGGCTTCCGATCCCTAAGGTAGGCCCTGGGCGCC 5404
DB 1139 ATGACATCCCGTGACATCGAAGTACAGGCTTCCGATCCCTAAGGTAGGCCCTGGGCGCC 1198
QY 5405 TCCTCACCCCGAGCTCAGCACACGACCTGTGTATAGCCCGCAGCATGCTACTGCCAGGTG 5464
DB 1199 TCCTCACCCCGAGCTCAGCACACGACCTGTGTATAGCCCGCAGCATGCTACTGCCAGGTG 1258
QY 5465 GGCCCACTCTAGGAACCTGGCCACCTAGTCTCAATGTCACCACTGACTGTCCCCAC 5524
```

RESULT 9

ABT33958

ID ABT33958 standard; DNA; 2170 BP.

XX

AC ABT33958;

XX DT 29-MAY-2003 (first entry)

XX DE Human pigmentation trait-related DNA - SEQ ID No 57.

XX Human; single nucleotide polymorphism; SNP; ds; melanocortin-1 receptor;
KW Genetic pigmentation trait; MC1R; agouti signaling protein; ASIP; race;
KW hair colour; eye colour; forensic tool.
XX Homo sapiens.
OS
XX WO200297047-A2.
XX
PD 05-DEC-2002.
XX
PF 28-MAY-2002; 2002WO-US016789.
XX
PR 25-MAY-2001; 2001US-0293560P.
PR 21-JUN-2001; 2001US-0300187P.
PR 07-AUG-2001; 2001US-0310781P.
PR 17-SEP-2001; 2001US-0323662P.
PR 26-OCT-2001; 2001US-0344418P.
PR 15-NOV-2001; 2001US-0334674P.
PR 02-JAN-2002; 2002US-0346303P.
XX
PA (DNAP-) DNAPRINT GENOMICS INC.
XX
XX Fridakis T;
XX WPI; 2003-239091/23.
XX
XX Inferring genetic pigmentation trait such as hair/eye color or shade from
PT nucleic acid sample of human subject, by identifying a pigmentation-
PT related haplotype allele of a pigmentation gene in the sample.
XX
XX Claim 50; Page 337-339; 396pp; English.
XX
XX The invention comprises a method for inferring a genetic pigmentation
CC trait of a human. The method involves identifying a single nucleotide
CC polymorphism (SNP) in a pigmentation gene - where the pigmentation gene
CC is not melanocortin-1 receptor (MC1R) and agouti signaling protein
CC (ASIP). The method of the invention is useful for inferring a genetic
CC pigmentation trait of a human, especially for inferring the race of a
CC human subject. The method is useful for inferring a genetic pigmentation
CC trait such as hair shade or colour, or eye shade or colour of a human
CC subject. The method may be used as a forensic tool for obtaining
CC information relating to physical characteristics of a potential crime
CC victim or a perpetrator of a crime from a nucleic acid sample present at
CC a crime scene. The present human DNA sequence is used in the
CC exemplification of the invention
XX
XX Sequence 2170 BP; 395 A; 650 C; 686 G; 424 T; 0 U; 15 Other;
XX
Query Match 19.5%; Score 1999.4; DB 7; Length 2170;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 2076; Conservative 14; Mismatches 66; Indels 16; Gaps 3;
XX 4219 GACATCTCAGACATGTCGTGGAGAGGTGTCGCCGGGTGACGGGSCACAGAGAGGCC 4278
DB 1 GACATCTCAGACATGTCGTGGAGAGGTGTCGCCGGGTGACGGGSCACAGAGAGGCC 60
XX 4279 AAGGACTCTGTACC-CCCGTCCACGTTGGAGATTTTCGATTTTAGGTTTCTCCTCTGGGCA 4337
DB 61 AAGGACTCTGTACCTCTATCAGGTGTCAGAGATTTTCGATTTTAGGTTTCTCCTCTGGGCA 120
XX 4338 AGGAGAGAGGTGGAGGCTGGGACTTTGGGAGGAGCTTGGTGAAGTGTAGTGTAGGACCA 4397
DB 121 AGGAGAGAGGTGGAGGCTGGGACTTTGGGAGGAGCTTGGTGAAGTGTAGTGTAGGACCA 180
XX 4398 GGCAGGCCCTGGTCTTACCTGGAGATGGCTGGGGCTGAGACTTGTCCAGGTTGAACGAG 4457
DB 181 GGCAGGCCCTGGTCTTACCTGGAGATGGCTGGGGCTGAGACTTGTCCAGGTTGAACGAG 240
XX 4458 AGCAGAGGAGGATTTGAGACCCCGTTCTGTCTGGTGTAGGTCGTAATCTGTCCCGGTC 4517
DB 241 AGCAGAGGAGGATTTGAGACCCCGTTCTGTCTGGTGTAGGTCGTAATCTGTCCCGGTC 300

QY 4518 CTCCTGCATCCACCGCGCTGCTGGCAAGGTCTCAAGCTTCCAAAAGGCTTTCTCTGACC 4577
DB 301 CTCCTGCATATCCACCGCGTGGCTGGCAAGGTCTCAAGCTTCCAAAAGGCTTTCTCTGACC 360
QY 4578 CAGCTGGATGAGCTGCTAACTCAGCACAGGATGACCTGGGACCCAGCCAGCCACCCCGA 4637
DB 361 CAGCTGGATGAGCTGCTAACTCAGCACAGGATGACCTGGGACCCAGCCAGCCACCCCGA 420
QY 4638 GACCTGACTGAGGCTTCTCTGGCAAGAAAGAGAGAGGTGAGAGTGGCTGCGACCGTGGG 4697
DB 421 GACCTGACTGAGGCTTCTCTGGCAGAGATGGAGAAAGGTGAGAGTGGCTGCGACCGTGGG 480
QY 4698 GGCAGGGGTGGGTGGGTGAAAGTCCACAGAGAAATGAGGGAGGCTGGGCAAAAGGTTGG 4757
DB 481 GGCAGGGGTGGGTGGGTGAGCGTCCAGAGAAATGAGGGAGGCTGGGCAAAAGGTTGG 540
QY 4758 ACCAGTGCATCACCCGGGAGCGCATCTGGGCTGACAGGTGCAGAAATTTGAGAGTCAITTT 4817
DB 541 ACCAGTGCATCACCCGGGAGCGCATCTGGGCTGACAGGTGCAGAAATTTGAGAGTCAITTT 600
QY 4818 GGGGGCTACCCGTTCTATCCCTGAGTATCTCTCGGCTCTGCTCAGGCCAAGGGAGC 4877
DB 601 GGGGGCTACCCGTTCTGTCCC--GAGTATGCTCTCGGCTCTGCTCAGGCCAAGGGAGC 658
QY 4878 CCTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGGGTAACTGTTCTCTGCTGCC 4937
DB 659 CCTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGGGTGAGTCTGTTCTCTGCTGCC 718
QY 4938 GGGAGTGTGACCACTCGACACAGCTGGCTGGGGCTCTCTGCTCATGATCTCTACACCTG 4997
DB 719 GGGAGTGTGACCACTCGACACAGCTGGCTGGGGCTCTCTGCTCATGATCTCTACATCG 778
QY 4998 GATGTCAGCTGAGCCAGCTGGGGCCCAAGGCAAGGACTGAGGAGGAAAGGGTACAGC 5057
DB 779 GATGTCAGCTGAGCCAGCTCTGGGAAACAGTGCAGGGGCCAGAGGAGGAGGTTACAGG 838
QY 5058 TGGGGGGCCCTGGGCTTGGTGGGACACCCGGGGCTTCCAGCACAGGCGTGGCCAGGCTC 5117
DB 839 CGGGGGCCCATGAATTTGCTGGGACACCCGGGGCTTCCAGCACAGGCTTGGACAGGATC 898
QY 5118 CTGTAAAGCTTAATCTCTCCAAACAGGAGGAAAGGAGAGTGT-----CCCT 5164
DB 899 CTGTAAAGCTTGAATCTCTCCAAACAGGAGGAAAGGAGTGTGAGGGCCGAGCCCT 958
QY 5165 GGGTCTGACCCATGTTGGGAGCATGTCGTTCAGTCCGTTGTCACAGAGATGAC 5224
DB 959 GGGTCTGACCCATGTTGGGAGCATGTCGTTCAGGCGGCTGTCTCCAGGAGGATGAC 1018
QY 5225 GACGTGATAGGGCAGGTGGGGCAGCAGAGATGGGTGACCGCTCAGATGCCCTACACC 5284
DB 1019 RACGTGATAGGGCAGGTGGGGCAGCAGAGATGGGTGACCGCTCAGATGCCCTACACC 1078
QY 5285 ACTGCGCTGATTTACAGAGGTGAGCGCTTTGGGACATCATCCCCCTGAGTGTGACCCAT 5344
DB 1079 ACTGCGCTGATTTCAAGAGTGCAGCGCTTTGGGACATCGTCCCTCGGTGTGACCCAT 1138
QY 5345 ATGACATCCGTCAGATCGAAAGTACAGGGCTTCCGATCCCTTAAGGTAGGCTTGGGGCC 5404
DB 1139 ATGACATCCGTCAGATCGAAAGTACAGGGCTTCCGATCCCTTAAGGTAGGCTTGGGGCC 1198
QY 5405 TCCTCACCCAGCTCAGCACCGGACCTGGGTGATAGCCCGCAGCATGGCTACTGCCAGGTG 5464
DB 1199 TCCTCACCCAGCTCAGCACCGGACCTGGGTGATAGCCCGCAGCATGGCTACTGCCAGGTG 1258
QY 5465 GGGCCACTCTAGGAACCTTGGCCACTAGTCTCTCAATGCCACCACTGACTGTCCCCAC 5524
DB 1259 GGGCCACTCTAGGAACCTTGGCCACTAGTCTCTCAATGCCACCACTGACTGTCCCCAC 1318
QY 5525 TTGGGTGGGGGTCCAGAGTATAGGAGGGCTGGCTGTCCATCCAGAGCCCCCGTCTAG 5584
DB 1319 TTGGGTGGGGGTCCAGAGTATAGGAGGGCTGGCTGTCCATCCAGAGCCCCCGTCTAG 1378
QY 5585 TGGGGAGACAAACAGGACCTGCCAGATTTGGAGGAGCCCGAGCCCTGCGAGGAGAGGG 5644


```
QY 4638 GACCTGACTGAGCCCTTCTGGCAAGAAGAGGATGAGAGTGGCTGCCACGCTGGGG 4697
Db 421 GACCTGACTGAGCCCTTCTGGCAGAGATGGAAGGTGAGTGGCTGCCACGCTGGGG 480
QY 4698 GGCAAGGGTGGTGGGTTGAACGTCCAGAGGAATGAGGGGAGGCTGGGCAAAAGGTTGG 4757
Db 481 GGCAAGGGTGGTGGGTTGAGCGTCCAGAGGAATGAGGGGAGGCTGGGCAAAAGGTTGG 540
QY 4758 ACCAGTGATCAACCCGGCAGCGCATCTGGGCTGACAGGTGAGATGAGATTTGAGAGTCAATTT 4817
Db 541 ACCAGTGATCAACCCGGCAGCGCATCTGGGCTGACAGGTGAGATTTGAGAGTCAATTT 600
QY 4818 GGGGGTACCCGTTCTATCCCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGGGAGC 4877
Db 601 GGGGGTACCCGTTCTGTFCCC--GAGTATGCTCTCGGCCCTGCTCAGGCCAAGGGGAAAC 658
QY 4878 CTTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGGTGGTAACTGTTCTCTTGCC 4937
Db 659 CTTGAGAGCAGCTTCAATGATGAGAACCTGCGCATAGTGGTGGTAACTGTTCTCTTGCC 718
QY 4938 GGGATGGTGAACACCTGCAACGCTGGCTGGGCTCTCTGCTCATGATCTACACCTG 4997
Db 719 GGGATGGTGAACACCTGCAACGCTGGCTGGGCTCTCTGCTCATGATCTACACCTG 778
QY 4998 GATGTGACGCTGAGCCAGCTGGGCCCAAGGCAGGACTGAGGAGGAAGGTACAGC 5057
Db 779 GATGTGACGCTGAGCCCATCTGGGAAACAGTGCAGGGGCCAGGAGGAAGGTACAGC 838
QY 5058 TGGGGGCCCTTGGGTTAGCTGGGCACACCCGGGGCTTCAGACACAGGCGTGGCCAGGCTC 5117
Db 839 CGGGGGCCCATGAACCTTTGCTGGGACACCCGGGGCTCCAAAGCACAGGCTTGAACAGGATC 898
QY 5118 CTGTAAGCCTAATCTCTCCAAACACAGAGGAAGGAGTGT-----CCCTC 5164
Db 899 CTGTAAGCCTGACCTCTCCAAATAGAGGAAGAAAGAGTGTGAGGGCCGAGCCCTC 958
QY 5165 GGGTGTGACCATTTGTGGGAGCATGTCGTGTCAGTCCGCTCCACAGAGATCGAC 5224
Db 959 GGGTGTGACCATTTGTGGGAGCATGTCGTGTCAGGCGGTGTCACAGGAGATCGAC 1018
QY 5225 GACGTGATAGGGCAGGTGGGGACCCAGAGATGGGTGACAGGCTCACATGCCCTACACC 5284
Db 1019 RACGTGATAGGCAGGTGGYGACAGAGATGGGTGACGCGCTCRATGCCCTRCAYC 1078
QY 5285 ACTGCCGTGATTCAGAGGTGAGCGCTTTGGGGACATCATCCCTCGATGTGACCCAT 5344
Db 1079 ACTGCCGTGATTCAGAGGTGAGCGCTTTGGGGACATCGTCCCTCGGTGACCCAT 1138
QY 5345 ATGATATCCCGTGACATCCAAGTACAGGGCTTCCGCTATCCCTAAGTAGGCTGGGCC 5404
Db 1139 ATGATATCCCGTGACATCCAAGTACAGGGCTTCCGCTATCCCTAAGTAGGCTGGGCC 1198
QY 5405 TCTCATCCCGCAGCTCAGCACAGCAGCTGGTGATAGCCCGCAGCATGGCTACTGCCAGGTG 5464
Db 1199 TCTCATCCCGCAGCTCAGCACAGCAGCTGGTGATAGCCCGCAGCATGGCTACTGCCAGGTG 1258
QY 5465 GGGCCACTCTAGAACCCCTGGCCACTAGTCTCTCAATGCCACCACTGACTGTGCCAC 5524
Db 1259 GGGCCACTCTAGAACCCCTGGCCACTAGTCTCTCAATGCCACCACTGACTGTGCCAC 1318
QY 5525 TTGGGTGGGGTCCAGAGTATAGGAGGGCTGGCTGTGTCATCCAGAGCCCGCTCTAG 5584
Db 1319 TTGGGTGGGGTCCAGAGTATAGGAGGGCTGGCTGTGTCATCCAGAGCCCGCTCTAG 1378
QY 5585 TGGGGAGACAAACCCAGGACCTGCCAGAAATTTGGAGGACCCAGCGCTGCGAGGAGGG 5644
Db 1379 TGGGGAGACAAACCCAGGACCTGCCAGAAATTTGGAGGACCCAGCGCTGCGAGGAGGG 1438
QY 5645 GGCAGTGTGGGTGCTCTGAGAGGTGTGATCTGGCCCTGCTGTGGGGTGGAGAGGGTAC 5704
Db 1439 GGCAGTGTGGGTGCTCTGAGAGGTGTGATCTGGCCCTGCTGTGGGGTGGAGAGGGTAC 1498
QY 5705 TGTGGAGCTTCTCGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCGCAGGAGTGTG 5764
```

RESULT 11

ADC26596

ID ADC26596 standard; DNA; 2170 BP.

XX AC ADC26596;

XX AC ADC26596;

XX DT 18-DEC-2003 (first entry)

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

XX DE Human CYP2D6 variant DNA - CYP2D657_286.

statin response; cytochrome p450 3A4; CYP3A4; 2D6; CYP2D6;
3-hydroxy-3-methylglutaryl-coenzyme A reductase; HMGCR; atorvastatin;
simvastatin; serum cholesterol level; heart attack;
single nucleotide polymorphism; SNP; human; ds.

Homo sapiens.

WO2003002721-A2.

09-JAN-2003.

01-JUL-2002; 2002WO-US020847.

29-JUN-2001; 2001US-0301867P.

07-AUG-2001; 2001US-0310783P.


```
QY 5885 TGAGGCGTCTGGGAGAGCCCTCGCTTCCACCCGACACACTTCTGGATGCCAGG 5944
DB 1679 TGAGGCGGCTGGGAGAGCCCTCGCTTCCACCCGACACACTTCTGGATGCCAGG 1738
QY 5945 CCACCTTTGTGAAGCCGAGGCGCTTCTGCTTTCTCAGCAGGTGCTGTGGGAGCCCGG 6004
DB 1739 CCACCTTTGTGAAGCCGAGGCGCTTCTGCTTTCTCAGCAGGTGCTGTGGGAGCCCGG 1798
QY 6005 CTCCTGTCTCCCTTCGCTGGAGTCTTGCAGGGGTATCACCCAGGAGCCAGGCTCACTGAC 6064
DB 1799 CTCCTGTCTCCCTTCGCTGGAGTCTTGCAGGGGTATCACCCAGGAGCCAGGCTCACTGAC 1858
QY 6065 GGCCTTCCCTCCACACAGGCGCGGTGATGCTGCGGGAGCCCTGCGCCGATGAG 6124
DB 1859 GGCCTTCCCTCCACACAGGCGCGGTGATGCTGCGGGAGCCCTGCGCCGATGAG 1918
QY 6125 CTCCTTCTCTCTTCACTCCCTGCTGTCAGACACTTTCAGCTTCTCGTGGCCGCGACAG 6184
DB 1919 CTCCTTCTCTCTTCACTCCCTGCTGTCAGACACTTTCAGCTTCTCGTGGCCGCGACAG 1978
QY 6185 CCCCGGCCAGGACACTCTCGTGTGCTGAGCTTCTTGTGACCCCATCCCTTACGAGCTT 6244
DB 1979 CCCCGGCCAGGACACTGCTGTCTTGTGCTTCTTGTGAGCCCATCCCTTATGAGCTT 2038
QY 6245 TGTGCTGTCCCGGTAGAAATGGGTACCTAGTCCCGAGCCTGCTCCCTAGCCAGAGCT 6304
DB 2039 TGTGCTGTCCCGGTAGAAATGGGTACCTAGTCCCGAGCCTGCTCCCTAGCCAGAGCT 2098
QY 6305 CTAATGTACAATAAGCAATGTGTAGTTTCCAACTTGGGTCCCTGCTCAAGCCCTCGTT 6364
DB 2099 CTAATGTACAATAAGCAATGTGTAGTTTCCAACTTGGGTCCCTGCTCAAGCCCTCGTT 2158
QY 6365 GGGATCATCCTC 6376
DB 2159 GGGATCATCCTC 2170

RESULT 12
ADC26591
ID ADC26591 standard; DNA; 2170 BP.
AC ADC26591;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human CYP2D6 variant DNA - CYP2D6E7_339.
XX
KW statin response; cytochrome p450 3A4; CYP3A4; 2D6; CYP2D6;
KW 3-hydroxy-3-methylglutaryl-coenzyme A reductase; HMGCR; atorvastatin;
KW simvastatin; serum cholesterol level; heart attack;
KW single nucleotide polymorphism; SNP; human; ds.
XX
OS Homo sapiens.
XX
XX WO2003002721-A2.
XX
XX 09-JAN-2003.
XX
XX 01-JUL-2002; 2002WO-US020847.
XX
XX 29-JUN-2001; 2001US-0301867P.
XX
XX 07-AUG-2001; 2001US-0310783P.
XX
XX 13-SEP-2001; 2001US-0322478P.
XX
XX (DNAP-) DNAPRINT GENOMICS INC.
XX
XX Prudakis T;
XX
XX WPI; 2003-239174/23.
XX
XX Inferring a statin response from a nucleic acid sample, by haplotype
XX allele indicative of statin response, a decrease in total cholesterol, or
```

PT in low density lipoprotein infers a statin response of the subject.

XX Claim 43; SEQ ID NO 1; 323pp; English.

CC The invention relates to a novel method for inferring a statin response from a nucleic acid sample comprising identifying in the nucleic acid sample, at least one haplotype allele indicative of a statin response. CC The haplotype allele may comprise nucleotides of the cytochrome p450 3A4 (CYP3A4) gene, nucleotides of the cytochrome p450 2D6 (CYP2D6) gene or nucleotides of the 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR) gene. The method of the invention may be useful for inferring a statin response of a human subject from a nucleic acid sample, where the human subject is a Caucasian subject and the statin is atorvastatin or simvastatin. The method may also be useful for determining whether to prescribe statin to a patient with elevated serum cholesterol levels in order to prevent heart attack. The current sequence is that of the human CYP2D6 variant DNA of the invention which contains a single nucleotide polymorphism.

XX Sequence 2170 BP; 395 A; 650 C; 686 G; 424 T; 0 U; 15 Other;

Query Match 19.5%; Score 1999.4; DB 9; Length 2170;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 2076; Conservative 14; Mismatches 66; Indels 16; Gaps 3;

QY 4219 GACATCTCAGACATGCTGCTGGAGAGGTGTCGCCGGTTCAGGGGGACACAGAGAGGCC 4278

DB 1 GACATCTCAGACATGCTGCTGGAGAGGTGTCGCCGGTTCAGGGGGACACAGAGAGGCC 60

QY 4279 AAGGACTCTGTACC-CCCGTCCACGTTGCGAGATTTCCGATTTTAGGTTTCTCTCTGGGCA 4337

DB 61 AAGGACTCTGTACCCTCTATCACTCAGAGATTTCCGATTTTAGGTTTCTCTCTGGGCA 120

QY 4338 AGGAGAGAGGTGGAGGCTGGCACCTTGGGGAGGGACTTTGGTAGGTCACTGTAAAGGACA 4397

DB 121 AGGAGAGAGGTGGAGGCTGGCACCTTGGGGAGGGACTTTGGTAGGTCACTGTAAAGGACA 180

QY 4398 GCGAGGCCCTGGGCTACTCTGAGATGGCTGGGGCTGAGACTTGTCCAGGTGAACGAG 4457

DB 181 GCGAGGCCCTGGGCTACTCTGAGATGGCTGGGGCTGAGACTTGTCCAGGTGAACGAG 240

QY 4458 AGCAGAGAGGATTTGAGACCCCGTTCTGTCTGGTGTAGGTGCTGAATGCTTCCCGCTC 4517

DB 241 AGCAGAGAGGATTTGAGACCCCGTTCTGTCTGGTGTAGGTGCTGAATGCTTCCCGCTC 300

QY 4518 CTCCTGCACATCCAGCGCTGGCTGGCAAGGTCTTAACGCTTCCAAAGGCTTTCCTGACC 4577

DB 301 CTCCTGCACATCCAGCGCTGGCTGGCAAGGTCTTAACGCTTCCAAAGGCTTTCCTGACC 360

QY 4578 CAGCTGGATGAGCTGCTTAACCTGAGCAGAGATGACCTGGGACCCAGCCAGCCAGCCGGA 4637

DB 361 CAGCTGGATGAGCTGCTTAACCTGAGCAGAGATGACCTGGGACCCAGCCAGCCAGCCGGA 420

QY 4638 GACCTGACTGAGGCTTCTCTGCGAAGAGAGAGGTGAGAGTGGCTGCCACGGTGGGG 4697

DB 421 GACCTGACTGAGGCTTCTCTGCGAAGAGAGAGGTGAGAGTGGCTGCCACGGTGGGG 480

QY 4698 GGCAGAGGTGTGGGTTCACCGTCCAGAGAGAAATGAGGGAGGCTGGGCAAAAGGTTGG 4757

DB 481 GGCAGAGGTGTGGGTTCACCGTCCAGAGAGAAATGAGGGAGGCTGGGCAAAAGGTTGG 540

QY 4758 ACCAGTGATCACCGGAGCGCATCTGGGCTGACAGAGTGCAGAAATGGAGGTCATTT 4817

DB 541 ACCAGTGATCACCGGAGCGCATCTGGGCTGACAGAGTGCAGAAATGGAGGTCATTT 600

QY 4818 GGGGGCTACCCGTTCTATCCCTGAGTATCTCTCGGCCCTGCTCAGGCCCAAGGGGAGC 4877

DB 601 GGGGGCTACCCGTTCTATCCCTGAGTATCTCTCGGCCCTGCTCAGGCCCAAGGGGAGC 658

QY 4878 CTTGAGAGCAGCTTCAATGATGAGAACCTGGCGCATAGTGGTGGGTAACTGTTCTTGGC 4937

DB 659 CTTGAGAGCAGCTTCAATGATGAGAACCTGGCGCATAGTGGTGGGTAACTGTTCTTGGC 718

```
QY 4938 GGGATGTTGACCACTCTGACCAAGCTGGCCCTGGGGCCCTCTCTGCTCATGATCTCTACACCTG 4997
DB 719 GGGATGTTGACCACTCTGACCAAGCTGGCCCTGGGGCCCTCTCTGCTCATGATCTCTACATCCG 778
QY 4998 GATGTGACAGCTGAGCCAGCTGGGGCCCAAGGACAGGAGTGTAGGGAGGAGGATACAGC 5057
DB 779 GATGTGACAGCTGAGCCAGCTCTGGGAAAGTGTAGGGGGCCGAGGGAGGAGGATACAGG 838
QY 5058 TGGGGGCCCCCTTGTAGCTGGGACACCCGGGGCTTCCAGCACAGGCGTGGCCAGGCTC 5117
DB 839 CGGGGGCCCATGAATTTGTCTGGGACACCCGGGGCTCCAGACACAGGCTTTCAGCAGGATC 898
QY 5118 CTGTAGACCTAACTTCTCTCAACACAGGAGGAGAGTGT-----CCCT 5164
DB 899 CTGTAGACCTGACCTCTCTCAACATATAGGAGGACAGGAGTGTTCAGGGCCGGACCCCT 958
QY 5165 GGGTGTGACCCCATTTGGGGAGCGATGTCTGTCCAGTCCGTGTCCAAACAGGAGATCGAC 5224
DB 959 GGGTGTGACCCCATTTGGGGAGCGRTGTCTGTCCAGGGCGTGTCCAAACAGGAGATCGAC 1018
QY 5225 GACGTGTATAGGGCAGGTGCGGGCAGCACAGAGATGGGTGACCAAGGCTCACATGCGCTACACC 5284
DB 1019 RACGTGTATAGGGCAGGTGCGGGCAGCACAGAGATGGGTGACCAAGGCTCACATGCGCTACACC 5284
QY 5285 ACTGCCGTGATTCACAGAGTGCAGCGCTTTTGGGACATCATCCCTCTGAGTGTGACCCAT 5344
DB 1079 ACTGCCGTGATTCACAGAGTGCAGCGCTTTTGGGACATCATCCCTCTGAGTGTGACCCAT 1138
QY 5345 ATGACATCCCGTGACATCGAAGTACAGGGCTTTCGGATCCCTAAGGTAGGCCCTGGGGCCC 5404
DB 1139 ATGACATCCCGTGACATCGAAGTACAGGGCTTTCGGATCCCTAAGGTAGGCCCTGGGGCCC 1198
QY 5405 TCCTCACCACCCAGCTCAGCACACAGCAGCTGTGTATAGCCCCAGCATGCTACTGCCAGGTG 5464
DB 1199 TCCTCACCACCCAGCTCAGCACACAGCAGCTGTGTATAGCCCCAGCATGCTACTGCCAGGTG 1258
QY 5465 GGCCCACTCTAGAAACCTGGCCACCCAGCTCTCAATGACCAACACACATGCTGCTCCACC 5524
DB 1259 GGCCCACTCTAGAAACCTGGCCACCCAGCTCTCAATGACCAACACACATGCTGCTCCACC 1318
QY 5525 TTGGTGGGGGTTCAGAGTATAGGACGGCTGGCTGTCTCATCAGAGCCCCGCTGTAG 5584
DB 1319 TTGGTGGGGGTTCAGAGTATAGGACGGCTGGCTGTCTCATCAGAGCCCCGCTGTAG 1378
QY 5585 TGGGAGACAAACACAGGACCTGCGAGAACTGTGGAGGACCCAGCGCTGCGAGGAGAGGG 5644
DB 1379 TGGGAGACAAACACAGGACCTGCGAGAACTGTGGAGGACCCAGCGCTGCGAGGAGAGGG 1438
QY 5645 GGCAGTGTGGGTGCTCTGAGAGGTGTGACTGGGCCCTGCTGTGGGTGCGAGAGGGTAC 5704
DB 1439 GGCAGTGTGGGTGCTCTGAGAGGTGTGACTGGGCCCTGCTGTGGGTGCGAGAGGGTAC 1498
QY 5705 TGTGGAGCTTCTCGGGCGAGGACTAGTTGACAGAGTCCAGCTGTGTGCGAGGACAGTGTG 5764
DB 1499 TGTGGAGCTTCTCGGGCGAGGACTAGTTGACAGAGTCCAGCTGTGTGCGAGGACAGTGTG 1558
QY 5765 TGTCCCGGTGTGTTGTGGAGGGGTCCAGCATCTCTAGAGTCCAGTCCCGCTCTCA 5824
DB 1559 TGTCCCGGTGTGTTGTGGAGGGGTCCAGCATCTCTAGAGTCCAGTCCCGCTCTCA 1618
QY 5825 CCCTGCATCTCTGCCAGGAGACGACATCTCATCAACCACTGCTGCTGCTGCTGAAGGA 5884
DB 1619 CCCTGCATCTCTGCCAGGAGACGACATCTCATCAACCACTGCTGCTGCTGCTGAAGGA 1678
QY 5885 TAGAGGCGTCTGGAGAGCCCTTTCGCTTCCACCCGGAACACTTCTCTGATGCCAGGG 5944
DB 1679 TAGAGGCGTCTGGAGAGCCCTTTCGCTTCCACCCGGAACACTTCTCTGATGCCAGGG 1738
QY 5945 CCACTTGTGAAGCCGAGGACCTTCTGCTTCTCAGCAGGTCCTGTGGGAGCCCG 6004
DB 1739 CCACTTGTGAAGCCGAGGACCTTCTGCTTCTCAGCAGGTCCTGTGGGAGCCCG 1798
QY 6005 CTCCTGTCCCTTCCGTGGAGTCTTGCAGGGGTATCACCCAGGAGCCAGGCTCACTGAC 6064

RESULT 13
ADC26595
ID ADC26595 standard; DNA; 2170 BP.
XX
AC ADC26595;
XX
DT 18-DEC-2003 (first entry)
XX
Human CYP2D6 variant DNA - CYP2D6E7_150.
XX
statin response; cytochrome p450 3A4; CYP3A4; 2D6; CYP2D6;
3-hydroxy-3-methylglutaryl-coenzyme A reductase; HMGCR; atorvastatin;
simvastatin; serum cholesterol level; heart attack;
single nucleotide polymorphism; SNP; human; ds.
XX
Homo sapiens.
XX
WO2003002721-A2.
XX
09-JAN-2003.
XX
01-JUL-2002; 2002WO-US020847.
XX
29-JUN-2001; 2001US-0301867P.
XX
07-AUG-2001; 2001US-0310783P.
XX
13-SEP-2001; 2001US-0322478P.
XX
(DNAP-) DNAPRINT GENOMICS INC.
XX
Prudakis T;
XX
WPI; 2003-239174/23.
XX
Inferring a statin response from a nucleic acid sample, by haplotype
allele indicative of statin response, a decrease in total cholesterol, or
in low density lipoprotein infers a statin response of the subject.
XX
Claim 162; SEQ ID NO 5; 323pp; English.
XX
The invention relates to a novel method for inferring a statin response
from a nucleic acid sample comprising identifying in the nucleic acid
sample, at least one haplotype allele indicative of a statin response.
XX
The haplotype allele may comprise nucleotides of the cytochrome p450 3A4
(CYP3A4) gene, nucleotides of the cytochrome p450 2D6 (CYP2D6) gene or
nucleotides of the 3-hydroxy-3-methylglutaryl-coenzyme A reductase
```


QY 6185 CCCGGCCAGCACCTCTGTCGTCAGCTTTCTGTCACCCCATCCCCCTACGAGCTT 6244
|||||
Db 1979 CCCGGCCAGCACCATGCTGCTTCTGTCAGCCCATCCCCCTATGAGCTT 2038
QY 6245 TGTGCTGTGCCCCGCTAGAAATGGGGTACCTAGTCCCAAGCTCTGCTCCCTAGCAGAGGCT 6304
Db 2039 TGTGCTGTGCCCCGCTAGAAATGGGGTACCTAGTCCCAAGCTCTGCTCCCTAGCAGAGGCT 2098
QY 6305 CTAATGTACAAATAAGCAATGCTGAGTTCCTCAACTTGGGTCCCTGCTCAGCCCTCGTT 6364
|||||
Db 2099 CTAATGTACAAATAAGCAATGCTGAGTTCCTCAACTTGGGTCCCTGCTCAGCCCTCGTT 2158
QY 6365 GGGATCATCCTC 6376
Db 2159 GGGATCATCCTC 2170

RESULT 14
ADC26792
ID ADC26792 standard; DNA; 2240 BP.
XX
AC ADC26792;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human lipitor/zocor response-related SNP DNA - SEQ ID 202.
XX
KW statin response; cytochrome p450 3A4; CYP3A4; 2D6; CYP2D6;
KW 3-hydroxy-3-methylglutaryl-coenzyme A reductase; HMGCR; atorvastatin;
KW simvastatin; serum cholesterol level; heart attack;
KW single nucleotide polymorphism; SNP; human; ds; lipitor; zocor.
XX
OS Homo sapiens.
XX
PN WO2003002721-A2.
XX
PD 09-JAN-2003.
XX
PF 01-JUL-2002; 2002WO-US020847.
XX
PR 29-JUN-2001; 2001US-0301867P.
PR 07-AUG-2001; 2001US-0310783P.
PR 13-SEP-2001; 2001US-0322478P.
XX
XX (DNAP-) DNAPRINT GENOMICS INC.
PA
XX
PI Fridakis T;
XX
XX WPI; 2003-239174/23.
XX
XX Inferring a statin response from a nucleic acid sample, by haplotype
PT allele indicative of statin response, a decrease in total cholesterol, or
PT in low density lipoprotein infers a statin response of the subject.
XX
XX
PS Example 9; SEQ ID NO 202; 323pp; English.
XX
XX The invention relates to a novel method for inferring a statin response
CC from a nucleic acid sample comprising identifying in the nucleic acid
CC sample, at least one haplotype allele indicative of a statin response.
CC The haplotype allele may comprise nucleotides of the cytochrome p450 3A4
CC (CYP3A4) gene, nucleotides of the cytochrome p450 2D6 (CYP2D6) gene or
CC nucleotides of the 3-hydroxy-3-methylglutaryl-coenzyme A reductase
CC (HMGCR) gene. The method of the invention may be useful for inferring a
CC statin response of a human subject from a nucleic acid sample, where the
CC human subject is a Caucasian subject and the statin is atorvastatin or
CC simvastatin. The method may also be useful for determining whether to
CC prescribe statin to a patient with elevated serum cholesterol levels in
CC order to prevent heart attack. The current sequence is that of the human
XX lipitor/zocor response-related SNP DNA of the invention.
XX
SQ Sequence 2240 BP; 466 A; 565 C; 771 G; 437 T; 0 U; 1 Other;

Query Match 19.3%; Score 1984.8; DB 9; Length 2240;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 2137; Conservative 1; Mismatches 78; Indels 36; Gaps 7;

QY 990 AATGTTTCCACAGATTTCTAATCAGAAACATGGAGCCAGAGCAGTGGAGAAATGACG 1049
|||||
Db 1 AAGGTTTCCACAGATTTCTAATCAGAAACATGGAGCCAGAAAGCAGTGGAGAGGACG 60
QY 1050 ACCCTCAGGAGCCCTGGAGGATGCTGTACAGGCTGGGGCAAGGGCTTTCAGGCTACCA 1109
|||||
Db 61 ACCCTCAGGAGCCCTGGAGGATGCTGTACAGGCTGGGGCAAGGGCTTTCAGGCTACCA 120
QY 1110 ACTGGAGCTCTGGGAACAGCCCTGTGCAAAACAGGAAGTCATGGCCCGCCAGAGCCCA 1169
|||||
Db 121 ACTGGAGCTCTGGGAACAGCCCTGTGCAAAACAGGAAGTCATGGCCCGCCAGAGCCCA 180
QY 1170 -GAAATGTGGCTGAGCTGGGATCCATGTACAGCTTTGAGGCTCACCGGAGCAGCTCT 1228
|||||
Db 181 GGAATGTGGCTGGCT-----GGGAGCAGCTCT 210
QY 1229 GGACAGGAGAGGTCCCATCCAGGAAACCTCGGSCATGGCTGGGAAGTGGGGTACTTGGTG 1288
|||||
Db 211 GGACAGGAGTGGTCCCATCCAGGAAACCTCCGGCATGGCTGGGAAGTGGGGTACTTGGTG 270
QY 1289 CCGGGTCTGTATGTGTGTGCTGCTGTGTGTGAGAGAGAAATGTGTGCCCTCAGTGTCA 1348
|||||
Db 271 CCGGGTCTGTATGTGTGTGCTGCTGTGTGTGAGAGAGAAATGTGTGCCCTCAGTGTCA 330
QY 1349 GTGTGAGTCTGTGTATGTGTGTAATATGTCTTTGTGTGGTGTATTTCTGCATGTGTAAAT 1408
|||||
Db 331 GTGTGAGTCTGTGTATGTGTGTAATATGTCTTTGTGTGGTGTATTTCTGCCTGTGTAAAT 390
QY 1409 CGTGTCCCTGCAAGTGTGAACAGTGTCTGGAGTGGGAGTGGCAAGAGATCTGTGC 1468
|||||
Db 391 CGTGTCCCTGCAAGTGTGAACAGTGTCTGGAGTGGGAGTGGCAAGAGATCTGTGC 450
QY 1469 ACCATCAGTGTGTGCATAGCGTCTGTGCATGTCAAGAGTCAAGGTGAAGTGAAGGAC 1528
|||||
Db 451 ACCATCAGTGTGTGCATAGCGTCTGTGCATGTCAAGAGTCAAGGTGAAGTGAAGGAC 510
QY 1529 CAGGCCATGATGCCACTCATCATCAGAGCTCTAAGGCCCCAGAGTGAAGTGCAGTGACA 1588
|||||
Db 511 CAGGCCATGATGCCACTCATCATCAGAGCTCTAAGGCCCCAGAGTGAAGTGCAGTGACA 570
QY 1589 GATAGGGTGTGAAGGTCACTCTGGAGTGGGAGTGGGGTAGGAAAGGCCAAGGTC 1648
|||||
Db 571 GATAGGGTGTGAAGGTCACTCTGGAGTGGGAGTGGGGTAGGAAAGGCCAAGGCC 630
QY 1649 ATGTTCTGGAGGAGGGGTGTGTGACTACTATTAGGGTGTATGAGCTAGCTGGGAGGTGGAT 1708
|||||
Db 631 ATGTTCTGGAGGAGGGGTGTGTGACTACTATTAGGGTGTATGAGCTAGCTGGGAGGTGGAT 690
QY 1709 GGC CGGGTCCACTGAGACCCCTGTTATCCAGAGCCTGTGTGGGCTTGGGGAGCTTGGGA 1768
|||||
Db 691 GGC CGGGTCCACTGAGAACCCCTGGTATCCAGAAAGGCTTTGCAGGCTTTCAGGAGCTTGA 750
QY 1769 GTGGGAGAGGGGGTGTGACTCTCCGACCAAGGCTTTCTACCACCTTACCTGGTGAAGGG 1828
|||||
Db 751 GTGGGAGAGGGGGTGTGACTCTCCGACCAAGGCTTTCTACCACCTTACCTGGTGAAGGG 810
QY 1829 CTGGAGCAGGAGGAGCAGCGGCAAGGACCTCTGGAGCAGGCCATACCTTGCCTGGCCTGAC 1888
|||||
Db 811 CTGGAGCAGGAGGAGCAGCGGCAAGGACCTCTGGAGCAGGCCATACCTTGCCTGGCCTGAC 870
QY 1889 TCTGCCACTTGGCAGCAGTCAACACAGCAGGTTCACTCAGCAGAGGCGCAAGGCCAT 1948
|||||
Db 871 TCTGCCACTTGGCAGCAGTCAACACAGCAGGTTCACTCAGCAGAGGCGCAAGGCCAT 930
QY 1949 CATCAGTCCCTTTATAAGGAGAGGTCAAGCGCTCGGTGTGCGGAGAGTGTCTCGCTG 2008
|||||
Db 931 CATCAGTCCCTTTATAAGGAGAGGTCAAGCGCTCGGTGTGCGGAGAGTGTCTCGCTG 990
QY 2009 GTCTCTGTGCTGTGGGTGGGGTGGCCAGGTGTGTCCAGAGGAGGCCCTTGGTAGT 2068

991 GTCTCTGTGCTGTGGGTGGGGTGCAGGTGTGTCCAGAGAGCCCAATTTGTAGT 1050
 2069 GAGGAGCCATGGGCTAGAACACTGGTGCCCTGGCCCATATAGTGGCCATCTTCGT 2128
 1051 GAGGAGAGTATGGGCTAGAACACTGGTGCCCTGGCCCATATAGTGGCCATCTTCGT 1110
 2129 CTCCTGGTGCATCTGATGACCGGACCAACGCTGGGCTGCACGCTTACCGCGAGTCCC 2188
 1111 CTCCTGGTGCATCTGATGACCGGACCAACGCTGGGCTGCACGCTTACCGCGAGTCCC 1170
 2189 CTCCTGCATCCCGGCTGGGCAACCTTGTGCTGATGTGGAATTTCCAGAACACCACTACTG 2248
 1171 CTCCTGCATCCCGGCTGGGCAACCTTGTGCTGATGTGGAATTTCCAGAACACCACTACTG 1229
 2249 CTCCTGCATCCCGGCTGGGCAACCTTGTGCTGATGTGGAATTTCCAGAACACCACTACTG 2308
 1230 CTCCTGCATCCCGGCTGGGCAACCTTGTGCTGATGTGGAATTTCCAGAACACCACTACTG 1289
 2309 ACCAGCAACATGGGTGGGTGTTAAACCAACAGGCTGGATCAGAGCCAGGCTGAGAAGG 2368
 1290 AGAAGCAACATGATGTGGTGGTGAACCAACAGGCTGGATCAGAGCCAGGCTGAGAAGG 1349
 2369 GGAAGCAGGTTTGGGGGACGTTCTCTGGGGAAGGACATTTATACATGGCATGAAGGACTGG 2428
 1350 GGAAGCAGGTTTGGGGGACG-TCTTGGAGAAGGCAATTTATACATGGCATGAAGGACTGG 1408
 2429 ATTTTCCAAAGCCAAAGGAGTAGGGCAAGGCTGGAGGTGGAGCTGGAGCTGGGCGAG 2488
 1409 ATTTTCCAAAGCCAAAGGAGTAGGGCAAGGCTGGAGGTGGAGCTGGAGCTGGGCGAG 1468
 2489 TGGGATGCAAGCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACA 2548
 1469 TGGGATGCAAGCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACA 1528
 2549 CCAGAGGAAGGCTTTGGGAATGGAAGTAGTGTAGTCTCTGAGTGCCCTTTAAATCAGG 2608
 1529 CCAGAGGAAGGCTTTGGGAATGGAAGTAGTGTAGTCTCTGAGTGCCCTTTAAATCAGG 1588
 2609 AAATCGAGGATGAAGGGGTGAGTGACCGGTTTAAACCTTTTGGACCTGGGGTCTCTCG 2668
 1589 AAATCGAGGATGAAGGGGTGAGTGACCGGTTTAAACCTTTTGGACCTGGGGTCTCTCG 1648
 2669 GGCCTCAGTG-CTCACCGGCATGGACCATCATCTGGGAATGGATGATTAACCTGGGGCTC 2727
 1649 GGCCTCAGTGCTCACCGGCATGGACCATCATCTGGGAATGGATGATTAACCTGGGGCTC 1708
 2728 TCGGCAATTTTGGTGACTTTGCAAGGTCTATACCTGGGTGAGCGCATCCAAACTGAGTTCC 2787
 1709 TCGGCAATTTTGGTGACTTTGCAAGGTCTATACCTGGGTGAGCGCATCCAAACTGAGTTCC 1768
 2788 TCATCACAAGAGGTGTGACCCCAACCCCTGCGCCCAACATCAGAGGCTGGGTCTCTCTCC 2847
 1769 TCATCACAAGAGGTGTGACCCCAACCCCTGCGCCCAACATCAGAGGCTGGGTCTCTCTCC 1828
 2848 TTCCACCTCTCACTCTGTAGCCCGGGGTCTGTCAGGTTCAATAGACTAGGAC 2907
 1829 TTCCACCTCTCACTCTGTAGCCCGGGGTCTGTCAGGTTCAATAGACTAGGAC 1888
 2908 CTGTAGTCTGGGCTGATCTGGCTTGACAAGAGGCTGACCCCTCTCTGAGTTGGG 2967
 1889 CTGTAGTCTGGGCTGATCTGGCTTGACAAGAGGCTGACCCCTCTCTGAGTTGGG 1948
 2968 CGCGCTTCGGGAGCGTGTTCAGCTGAGCTGGGCTGAGCGCGGTGTCTGTCTCAAT 3027
 1949 CGCGCTTCGGGAGCGTGTTCAGCTGAGCTGGGCTGAGCGCGGTGTCTGTCTCAAT 2008
 3028 GGGCTGGCGGCTGCGGAGCGGATGTGACCGGGGAGGACACCGCCAGCCCGCG 3087
 2009 GGGCTGGCGGCTGCGGAGCGGCTGTGTGACCGGGGAGGACACCGCCAGCCCGCG 2068
 3088 CTGTGGCCCATCTACAGGCTCTGGGCTTCTGGGCGCGCTTCCAAAGGCAAGCGGGTGG 3147

Db 2069 CCTGTGCCCATCACCATGATCCTGGGTTTCGGGCGCGCTTCCAAAGCAAGCAGCGGT-G 2127
 QY 3148 GGGACAGAGACCGCTTTCCGTGGGCGCCCGGTGACAGTACCTAGCCCAAGCAGCGC 3207
 Db 2128 GGGACAGAGACAG-ATTTCCGTGGGACCCCGGTGGGTATGACCTAGTCCGAGCTGGGC 2186
 QY 3208 CCACAGGCGGTGGGGTCTCTGGACGTGAAACAG 3239
 Db 2187 AGAGAGGGCGCGGCTCTGGACATGAACAG 2218

RESULT 15
ABT33960

ID ABT33960 standard; DNA; 2240 BP.

XX ABT33960;

XX 29-MAY-2003 (first entry)

XX Human pigmentation trait-related DNA - SEQ ID No 59.

XX Human; single nucleotide polymorphism; SNP; ds; melanocortin-1 receptor;
 KW genetic pigmentation trait; MC1R; agouti signaling protein; ASIP; race;
 KW hair colour; eye colour; forensic tool.

XX Homo sapiens.

XX WO200297047-A2.

XX 05-DEC-2002.

XX 28-MAY-2002; 2002WO-US016789.

XX 25-MAY-2001; 2001US-0293560P.

XX 21-JUN-2001; 2001US-0300187P.

XX 07-AUG-2001; 2001US-0310781P.

XX 17-SEP-2001; 2001US-0323662P.

XX 26-OCT-2001; 2001US-0344418P.

XX 15-NOV-2001; 2001US-0334674P.

XX 02-JAN-2002; 2002US-0346303P.

XX (DNAP-) DNAPRINT GENOMICS INC.

XX Prudakis T;

XX WPI; 2003-239091/23.

XX Claim 50; Page 340-341; 396pp; English.

The invention comprises a method for inferring a genetic pigmentation trait of a human. The method involves identifying a single nucleotide polymorphism (SNP) in a pigmentation gene - where the pigmentation gene is not melanocortin-1 receptor (MC1R) and agouti signaling protein (ASIP). The method of the invention is useful for inferring a genetic pigmentation trait of a human, especially for inferring the race of a human subject. The method is useful for inferring a genetic pigmentation trait such as hair shade or colour, or eye shade or colour of a human subject. The method may be used as a forensic tool for obtaining information relating to physical characteristics of a potential crime victim or a perpetrator of a crime from a nucleic acid sample present at a crime scene. The present human DNA sequence is used in the exemplification of the invention

XX SQ Sequence 2240 BP; 466 A; 565 C; 771 G; 437 T; 0 U; 1 Other;

Query Match 19.3%; Score 1984.2; DB 7; Length 2240;
 Best Local Similarity 94.9%; Pred. No. 0;
 Matches 2137; Conservative 0; Mismatches 79; Indels 36; Gaps 7;

QY 990 AATGTTTCCACAGATTCTTAATCAGAAACATGAGGCCAGGAAGCATGTGGAGATGACG 1049
Db 1 AACCTTCCCAACAGATTCTTAATCAGAAACATGAGGCCAGGAAGCATGTGGAGAGACG 60
QY 1050 ACCCTCAGGAGCCCTGAGAGATGCTGACAGGCTGGGGCAAGGGCTTTCAAGCTACCA 1109
Db 61 ACCCTCAGGAGCCCTGAGAGATGCTGACAGGCTGGGGCAAGGGCTTTCCGGCTACCA 120
QY 1110 ACTGGAGCTCTGGGAACAGCCCTGTTGCAAAACAGGAAGTCAATGGCCCGGCCAGAGCCCA 1169
Db 121 ACTGGAGCTCTGGGAACAGCCCTGTTGCAAAACAGGAAGCCATAGCCCGGCCAGAGCCCA 180
QY 1170 -GAATGAGCTGAGCTGGGATCCATGTGACAGCTTTGAGGCTCACGGGAGAGAGCTCT 1228
Db 181 GGAATGAGGCTGGGCT-----GGGAGAGAGCTCT 210
QY 1229 GGAAGAGAGGCTCCCATCCAGGAACCTCGGCGATGGCTGGGAAGTGGGGTACTTTGGTG 1288
Db 211 GGAAGAGAGGCTCCCATCCAGGAACCTCGGCGATGGCTGGGAAGTGGGGTACTTTGGTG 270
QY 1289 CCGGGTCTGTATGTGTGTGACTGGTGTGTGTGAGAGAAATGTGTGCCCTGAGTGTCA 1348
Db 271 CCGGGTCTGTATGTGTGTGACTGGTGTGTGTGAGAGAAATGTGTGCCCTAAGTGTCA 330
QY 1349 GTGTGAGTCTGTGTGTGTAATATTTCTTTGTGTGGTGAATTTCTGCATGTGTAAAT 1408
Db 331 GTGTGAGTCTGTGTGTGTAATATTTCTTTGTGTGGTGAATTTCTGCCTGTGTAAAT 390
QY 1409 CGTGTCCCTCAGTGTGAACAAGTGGACAAGTGTCTGGAGTGGACAAGAGATCTGTGC 1468
Db 391 CGTGTCCCTCAGTGTGAACAAGTGGACAAGTGTCTGGAGTGGACAAGAGATCTGTGC 450
QY 1469 ACCATCAGTGTGTGATAGCTCTGTGCAATGTCAAGAGTGCAGAGTGAAGTCAAGGGAC 1528
Db 451 ACCATCAGTGTGTGATAGCTCTGTGCAATGTCAAGAGTGCAGAGTGAAGTCAAGGGAC 510
QY 1529 CAGGCCATGATGCCATCATCATCAGGAGCTTAAGGCCCAAGTAAGTGCAGTGAACA 1588
Db 511 CAGGCCATGATGCCATCATCATCAGGAGCTCTAAGGCCCAAGTAAGTGCAGTGAACA 570
QY 1589 GATAAGGCTCTCAAGTCTACTCGAGTGGGAGGCTGGGGTAGGGAAGGCAAGGTC 1648
Db 571 GATAAGGCTCTCAAGTCTACTCGAGTGGGAGGCTGGGGTAGGGAAGGCAAGGTC 630
QY 1649 ATGTTCTGAGGAGGGGTGTGACTACATTAAGGCTGTATGAGCTAGCTGGAGGTGGAT 1708
Db 631 ATGTTCTGAGGAGGGGTGTGACTACATTAAGGCTGTATGAGCTAGCTGGAGGTGGAT 690
QY 1709 GCGCGGTCTCATGAGACCTGTGTTATCCAGAGGCTGTGTGGGCTTGGGGAGCTTGA 1768
Db 691 GCGCGGTCTCATGAAACCTGTGTTATCCAGAGGCTTTGACAGGCTTCAGGAGCTTGA 750
QY 1769 GTGGGAGAGGGGTGACTTCTCCGACCAAGGCTTTTACCACTACCTACCTGGGTAAAGG 1828
Db 751 GTGGGAGAGGGGTGACTTCTCCGACCAAGGCTTTTACCACTACCTACCTGGGTAAAGG 810
QY 1829 CTTGAGAGGAGGAGCAGCGGCAAGGACCTCTGAGAGGCTTACCTGCTGGCTGAC 1888
Db 811 CTTGAGAGGAGGAGCAGCGGCAAGGACCTCTGAGAGGCTTACCTGCTGGCTGAC 870
QY 1889 TCTGCCACTGGCAGCAAGTCAACAGCAGAGTTCATCAGCAGAGGCGAGAGGCCAT 1948
Db 871 TCTGCCACTGGCAGCAAGTCAACAGCAGAGTTCATCAGCAGAGGCGAGAGGCCAT 930
QY 1949 CATCAGTCCCTTTATAGGAGAGGCTCAGCGCTGTGTGTGCGAGAGTGTCTTGCCTG 2008
Db 931 CATCAGTCCCTTTATAGGAGAGGCTCAGCGCTGTGTGTGCGAGAGTGTCTTGCCTG 990
QY 2009 GTCTCTGTGCTGTGGGTGGGGGTGCGAGGTGTCCAGAGAGGCCAGTTGGTAGT 2068
Db 991 GTCTCTGTGCTGTGGGTGGGGGTGCGAGGTGTCCAGAGAGGCCATTTGGTAGT 1050
QY 2069 GAGGAGGCCATGGGGCTAGAAAGCACTGGTGCCCTGGCCATGATAGTGGCCATCTTCCCTG 2128

Db 1051 GAGGAGGATATGGGGCTAGAAAGCACTGGTGCCCTGGCGTGTAGTGGGCATCTTCCCTG 1110
QY 2129 CTCCTGTGTGAGACCTGATGCAACCGGCACCAACGCTGGGCTGCAAGCTACCGCCAGGTCCC 2188
Db 1111 CTCCTGTGTGAGACCTGATGCAACCGGCACCAACGCTGGGCTGCAAGCTACCGCCAGGCCCC 1170
QY 2189 CTGCCACTGCCCCGGGCTGGGCAACCTTGTGCTGCAATGTGGACTTCCAGAAACACCATACTG 2248
Db 1171 CTGCCACTGCCCCGGGCTGGGCAACCTTGTGCTGCAATGTGGACTTCCAGAAACACCATACTG 1229
QY 2249 CTTTGCACAGGTGAGGAGGAGGCTCTTGGAGGGCGGAGAGGCTCTTGAAGATGCCCCACC 2308
Db 1230 CTTTGCACAGGTGAGGAGGAGGCTCTTGGAGGGCGGAGAGGCTCTTGAAGATGCCCCACC 1289
QY 2309 ACCAGCAAAACATGCGGTGGGTGAAACCAACAGGCTGGATCAGAAGCAGGCTGAGAAGG 2368
Db 1290 AGNAGCAAAACATGATGCTGGGTGAAACCAACAGGCTGGACAGNAGCGGCTGAGAAGG 1349
QY 2369 GGAAGCAGGTTTGGGGGACCTTCTCGGGGAAGGACATTTATATCATGGCATGAAGGACTGG 2428
Db 1350 GGAAGCAGGTTTGGGGGACG-TCCTGGAGAAAGGCAATTTATATCATGGCATGAAGGACTGG 1408
QY 2429 ATTTTCCAAAGGCCAAGGAAGTAGGCAAGGGCTTGGAGGTGGAGCTTGGCAG 2488
Db 1409 ATTTTCCAAAGGCCAAGGAAGTAGGCAAGGGCTTGGAGGTGGAGCTTGGCAG 1468
QY 2489 TGGGATGCAAGCCCATTTGGGCAACATATGTTATGAGGTACAAAGTCCCTTCTGCTGACA 2548
Db 1469 TGGGATGCAAGCCCATTTGGGCAACATATGTTATGAGGTACAAAGTCCCTTCTGCTGACA 1528
QY 2549 CCAGAGAAAGGCTTTGGGAATGGAAGATGATTTAGTCTCTGAGTGCCTTTAAATCAGC 2608
Db 1529 CCAGAGAAAGGCTTTGGGAATGGAAGATGATTTAGTCTCTGAGTGCCTTTAAATCAGC 1588
QY 2609 AAATCAGAGATGAAGGGGTGAGTGAACCCGGTTCAAACTTTTGTGCACTGTGGGTCTCG 2668
Db 1589 AAATCAGAGATGAAGGGGTGAGTGAACCCGGTTCAAACTTTTGTGCACTGTGGGTCTCG 1648
QY 2669 GGCCTCAGCT-CTCAGCCGATGGACCATCATCTGGGAATGGATGCTAACTGGGGCTC 2727
Db 1649 GGCCTCAGCTCAGCCGATGGACCATCATCTGGGAATGGATGCTAACTGGGGCTC 1708
QY 2728 TCGGCAATTTTGTGACTCTTGCAGGTCATCTGGGTGACGATCCAACTGAGTTC 2787
Db 1709 TCGGCAATTTTGTGACTCTTGCAGGTCATCTGGGTGACGATCCAACTGAGTTC 1768
QY 2788 TCCATCAGAGAGGTGACCCCCACCCCTGCCCCACGATCAGAGGCTGGGTCTCTCTC 2847
Db 1769 TCCATCAGAGAGGTGACCCCCACCCCTGCCCCACGATCAGAGGCTGGGTCTCTCTC 1828
QY 2848 TTTCCACTCTCCTCTGCTAGCCCGGGGTGCTTCAAGGTTCAAAATAGGACTAGGAC 2907
Db 1829 TTTCCACTCTCCTCTGCTAGCCCGGGGTGCTTCAAGGTTCAAAATAGGACTAGGAC 1888
QY 2908 CTGTAGTCTGGGTGATCTCTGGCTTGAAGAGGCTTGAACCTCTCTGCAAGTTCGCG 2967
Db 1889 CTGTAGTCTGGGTGATCTCTGGCTTGAAGAGGCTTGAACCTCTCTGCAAGTTCGCG 1948
QY 2968 CGCGCTTCCGGGAAGTGTTCAGCTGAGCTGGCTGAGCGCGGTGGTCTGTCTCAAT 3027
Db 1949 CGCGCTTCCGGGAAGTGTTCAGCTGAGCTGGCTGAGCGCGGTGGTCTGTCTCAAT 2008
QY 3028 GGGCTGGCGGCTGCGGAGGCGATGTTGACCCCGGCGAGGACAGCGCCGACCCCG 3087
Db 2009 GGGCTGGCGGCTGCGGAGGCGATGTTGACCCCGGCGAGGACAGCGCCGACCCCG 2068
QY 3088 CTTGCGCCCATCTACAGGCTCTGGGCTTGGGCTTCCAAAGGCAAGCGCGGTGG 3147
Db 2069 CTTGCGCCCATCTACAGGCTCTGGGCTTGGGCTTCCAAAGGCAAGCGCGGTGG- 2127
QY 3148 GGGCAGAGACCGGCTTCCGTGGGCTTGGAGTGGAGCGGTAGCGGTAGCGCAAGCAGCGC 3207

Db 2128 GGGACAGACAG-ATTTCCGTGGGACCCGGGTGGGTGATGACCGTAGTCCGAGCTGGGC 2186
Qy 3208 CGACAGGGCGTGGGTCTCTGGACGTGAACAG 3239
Db 2187 AGAGAGGGCGCGGGTCTGTGGACATGAAACAG 2218

Search completed: February 28, 2004, 03:29:25
Job time : 2404.9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 28, 2004, 01:41:15 ; Search time 25578.9 Seconds
(without alignments)
17415.928 Million cell updates/sec
Title: US-09-820-788A-3
Perfect score: 10278
Sequence: 1 agcctacaaagtctggga.....ccagggtcagtcggcaggt 10278

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 segs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	8254.6	80.3	114846	9	HS257120	AL021878 Human DNA
C 2	8173.8	79.5	133246	9	BX247885	BX247885 Human DNA
C 3	7169.4	69.8	13677	9	HSCYP2D7B	X58468 Human CYP2D
C 4	6908	67.2	13278	9	HSCYP2D7A	X58467 Human CYP2D
C 5	6413.4	62.4	17060	9	HUMCYP2P	M3387 Human debri
C 6	5325.6	51.8	9432	6	AX394456	AX394456 Sequence
C 7	5325.6	51.8	9432	6	AX687027	AX687027 Sequence
C 8	5325.6	51.8	9432	6	HUMCYP2D6	M3388 Human cyloc
C 9	5314.6	51.7	9433	6	AX687028	AX687028 Sequence
C 10	4854	47.2	5503	9	HUMCYP2DG	M3389 Human debri
C 11	1927.6	18.8	5884	6	AX345458	AX345458 Sequence
C 12	1927.6	18.8	5884	6	AX348344	AX348344 Sequence
C 13	1631.6	15.9	5884	6	AX345459	AX345459 Sequence
C 14	1631.6	15.9	5884	6	AX348345	AX348345 Sequence
C 15	1330	12.9	1450	6	AX192411	AX192411 Sequence
C 16	1220.4	11.9	1680	6	AX394457	AX394457 Sequence
C 17	1207.4	11.7	1669	6	AX207224	AX207224 Sequence
C 18	1068.8	10.4	1506	9	BC037807	BC037807 Homo sapi
C 19	940.4	9.1	179691	4	BX296515	BX296515 Pig DNA s
C 20	877.2	8.5	179798	9	AL359532	AL359532 Human DNA
C 21	875.4	8.5	173509	9	AP002436	AP002436 Homo sapi
C 22	875.4	8.5	176053	2	AC068190	AC068190 Homo sapi
C 23	875.4	8.5	187946	9	AP003402	AP003402 Homo sapi
C 24	875.4	8.5	190960	2	AC036188	AC036188 Homo sapi
C 25	873.2	8.5	185624	9	AC103923	AC103923 Homo sapi
C 26	872.6	8.5	187204	9	AC084082	AC084082 Homo sapi
C 27	866.8	8.4	1315	9	HS245011	X16866 Human mRNA
C 28	862.6	8.4	99592	9	AF263284	AF263284 Homo sapi
C 29	862.6	8.4	191231	9	AC131005	AC131005 Homo sapi
C 30	861.6	8.4	330836	2	AC108703	AC108703 Homo sapi
C 31	844.6	8.2	77111	9	AC097655	AC097655 Homo sapi
C 32	841.4	8.2	185228	9	CNS01RGU	AL160192 Human chr
C 33	837.8	8.2	55284	9	AL136318	AL136318 Human DNA
C 34	831.4	8.1	134060	9	AC010389	AC010389 Homo sapi
C 35	829.8	8.1	163673	9	AC093240	AC093240 Homo sapi
C 36	827.8	8.1	157152	9	AC073427	AC073427 Homo sapi
C 37	824.6	8.0	173821	9	AC020599	AC020599 Homo sapi
C 38	813.8	7.9	159667	9	AC078953	AC078953 Homo sapi
C 39	813.2	7.9	142494	9	AC115112	AC115112 Homo sapi
C 40	810.6	7.9	164194	9	AC108477	AC108477 Homo sapi
C 41	809.8	7.9	191134	9	AC006065	AC006065 Homo sapi
C 42	809.2	7.9	165067	9	AL138927	AL138927 Human DNA
C 43	805.2	7.8	166381	2	AC116943	AC116943 Pan trogl
C 44	804.2	7.8	125977	9	CNS057C5	AL355053 Human chr
C 45	803.6	7.8	211515	9	AC144479	AC144479 Pan trogl

ALIGNMENTS

RESULT 1
HS257120/c
LOCUS
DEFINITION Human DNA sequence from clone Rpl-257120 on chromosome 22q13.1-13.2, complete sequence.
ACCESSION AL021878
VERSION AL021878.2 GI:17065905
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 114846)
TITLE Bridgeman,A.
Direct Submission

JOURNAL

Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

COMMENT

On Nov 23, 2004 this sequence version replaced G1240112. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

```
repeat_region 66514..66549
misc_feature /note="18 copies 2 mer aa 80% conserved"
66530..66802
misc_feature /note="match: GSS: Em:AQ617313"
66577..66930
misc_feature /note="match: GSS: Em:AQ084812"
66603..66893
misc_feature /note="match: GSS: Em:AQ044403"
66974..67017
repeat_region /note="22 copies 2 mer cc 75% conserved"
misc_feature complement(69196..69374)
misc_feature /note="match: GSS: Em:B14383"
69710..70103
misc_feature /note="match: GSS: Em:AQ181535"
misc_feature complement(78378..78961)
misc_feature /note="match: GSS: Em:AQ389013"
misc_feature complement(78421..78533)
misc_feature /note="match: GSS: Em:AQ042556"
misc_feature complement(78451..78977)
misc_feature /note="match: GSS: Em:AQ533248"
misc_feature complement(81846..82274)
misc_feature /note="match: GSS: Em:AQ618257"
82587..83003
repeat_region /note="L1M5A repeat: matches 5857..6292 of consensus"
83609..83702
repeat_region /note="MER21B repeat: matches 703..787 of consensus"
83697..83797
repeat_region /note="MER21B repeat: matches 686..794 of consensus"
83910..84357
repeat_region /note="MER21B repeat: matches 76..540 of consensus"
85375..85430
repeat_region /note="28 copies 2 mer aa 76% conserved"
misc_feature complement(90629..90900)
repeat_region /note="match: GSS: Em:B13983"
96343..96364
misc_feature /note="11 copies 2 mer ac 100% conserved"
99274..99444
repeat_region /note="Other . Weak data"
104028..104091
misc_feature /note="32 copies 2 mer aa 67% conserved"
complement(107923..108378)

Query Match 80.3%; Score 8254.6; DB 9; Length 114846;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 8319; Conservative 0; Mismatches 14; Indels 6; Gaps 5;

QY 1 AGCCTTACAAAGTCTGGGATTAACCTGCGTGAGCCACCGGTCGGCCTCTTTATCTCTT 60
DB 52430 AGCCTTACAAAGTCTGGGATTAACCTGCGTGAGCCACCGGTCGGCCTCTTTATCTCTT 52371

QY 61 ACTGTACTGTCTCTTTGAAAGTACTTATTTATTTTGTGTTTCATCATTTAGTCTAA 120
DB 52370 ACTGTACTGTCTCTTTGAAAGTACTTATTTATTTTGTGTTTCATCATTTAGTCTAA 52311

QY 121 TTAATAAGAGTAGTTTACACCAATTAAGATTAATAATCTCTGTTTCTGTG 180
DB 52310 TTAATAAGAGTAGTTTACACCAATTAAGATTAATAATCTCTGTTTCTGTG 52251

QY 181 TGCCTTACTATTACAGTGAGTTTGTACCTTTAGATGATTTCTTCTGCTCATTATATC 240
DB 52250 TGCCTTACTATTACAGTGAGTTTGTACCTTTAGATGATTTCTTCTGCTCATTATATC 52191

QY 241 CTTTCTTTTTCAGATTGAAAACTCCCTTTAGCATTTCTTGTGGGATATAGTCTGTTG 300
DB 52190 CTTTCTTTTTCAGATTGAAAACTCCCTTTAGCATTTCTTGTGGGATATAGTCTGTTG 52131

QY 301 TTGATGAATCTCGCAGCTTTTGTCTGTGGAGGTCCTTTATTTCTCCTTCTGTG 360
DB 52130 TTGATGAATCTCGCAGCTTTTGTCTGTGGAGGTCCTTTATTTCTCCTTCTGTG 52071

QY 361 AAGGATATTTTCCGAGATAGTTTATCTAGGCTAAAAGTTTCTTCCCTCAGCACTTT 420
DB 52070 AAGGATATTTTCCGAGATAGTTTATCTAGGCTAAAAGTTTCTTCCCTCAGCACTTT 52011
```

```
QY 421 AAATATGTATGCCACTCCCCCTGCTGTAAAGTTTCCACTGGAAGTGTGCTGCCCC 480
DB 52010 AAATATGTATGCCACTCCCCCTGCTGTAAAGTTTCCACTGGAAGTGTGCTGCCCC 51951

QY 481 ATGTATGTATTTGGAGCTCTACTGCATGTATTTGTTCTTTCTCTGCTGCTTTAGG 540
DB 51950 ATGTATGTATTTGGAGCTCTACTGCATGTATTTGTTCTTTCTCTGCTGCTTTAGG 51891

QY 541 ATCTCTTTCTTTATCTCTTGGACCTTTTCGAGTGTAAATATCAGATGCCCTTGAAGTCTGCTTC 600
DB 51890 ATCTCTTTCTTTATCTCTTGGACCTTTTCGAGTGTAAATATCAGATGCCCTTGAAGTCTGCTTC 51831

QY 601 TTTGGTTAAATCTGCTGCTGCTTCTATAAATCTTTGTTACAAAAATCAGCAGGATG 660
DB 51830 TTTGGTTAAATCTGCTGCTGCTTCTATAAATCTTTGTTACAAAAATCAGCAGGATG 51771

QY 661 GTGGTGGGACCTGTAAATCCAGCTACTTTGGAGGCTGAGGAGGAGATCGCTTGAACC 720
DB 51770 GTGGTGGGACCTGTAAATCCAGCTACTTTGGAGGCTGAGGAGGAGATCGCTTGAACC 51711

QY 721 CTGGAGGTGGAGTTCAGCTGAGCCGAGATCGCATCATTTGCACTCCACCTGGGCGACAG 780
DB 51710 CTGGAGGTGGAGTTCAGCTGAGCCGAGATCGCATCATTTGCACTCCACCTGGGCGACAG 51651

QY 781 AGCAAAATCCGTCTCAAAAAAATTTATTTGGGCTCGGTGCTGTGTAGTCCAGCT 840
DB 51650 AGCAAAATCCGTCTCAAAAAAATTTATTTGGGCTCGGTGCTGTGTAGTCCAGCT 51591

QY 841 ACTTGGGAGGAGGAGTCCACTTGATGTGAGATGCGATGAGGAGGATGATCTGCACT 900
DB 51590 ACTTGGGAGGAGGAGTCCACTTGATGTGAGATGCGATGAGGAGGATGATCTGCACT 51531

QY 901 GCACTCCGCGCGGCAACAGAGTGAGACCTGTCTTAAAGAAAAATAAAAAAAG 960
DB 51530 GCACTCCGCGCGGCAACAGAGTGAGACCTGTCTTAAAGAAAAATAAAAAAAG 51471

QY 961 CAACATATCTTAAATAAAGATCCTCCATAATGTTTCCACCAAGATTTCTAATCAGAAACA 1020
DB 51470 CAACATATCTTAAATAAAGATCCTCCATAATGTTTCCACCAAGATTTCTAATCAGAAACA 51411

QY 1021 TGGAGCCAGGAAGAGTGAGGAGATGACACCCCTCAGGAGAGGCTGAGAGATGCTGTAC 1080
DB 51410 TGGAGCCAGGAAGAGTGAGGAGATGACACCCCTCAGGAGAGGCTGAGAGATGCTGTAC 51351

QY 1081 AGGCTGGGCAAGGGGCTTTGAGCTACCACTGGGAGCTCTGGGAGAGCCCTGTTGCAA 1140
DB 51350 AGGCTGGGCAAGGGGCTTTGAGCTACCACTGGGAGCTCTGGGAGAGCCCTGTTGCAA 51291

QY 1141 ACAGGAAGTCAATGGCCCGGCGGAGAGCCAGAAATGTGGGCTGAGCTGGGATCCAATGTGACA 1200
DB 51290 ACAGGAAGTCAATGGCCCGGCGGAGAGCCAGAAATGTGGGCTGAGCTGGGATCCAATGTGACA 51231

QY 1201 GCTTTGAGGCTCACCGGAGAGGCTCTGAGACAGGAGGTCCTATCCAGGAACCTCGG 1260
DB 51230 GCTTTGAGGCTCACCGGAGAGGCTCTGAGACAGGAGGTCCTATCCAGGAACCTCGG 51171

QY 1261 GCATGCTGGGAGTGGGCTTCTGTCGCGGCTCTGTATGTGTGCTGCTGCTGCTG 1320
DB 51170 GCATGCTGGGAGTGGGCTTCTGTCGCGGCTCTGTATGTGTGCTGCTGCTGCTG 51111

QY 1321 TGAGAGAGATGTGTGCCCTGAGTGTGAGTGTGATGTGTGATGTGTGAATTTGCTT 1380
DB 51110 TGAGAGAGATGTGTGCCCTGAGTGTGAGTGTGATGTGTGATGTGTGAATTTGCTT 51051

QY 1381 TGTGTGGTGTATTTCTGCATGTGTAATCGTGTCCCTCAAGTGTGAAACAGTGGACAAG 1440
DB 51050 TGTGTGGTGTATTTCTGCATGTGTAATCGTGTCCCTCAAGTGTGAAACAGTGGACAAG 50991

QY 1441 TGTCTGGGAGTGGACAAGAGATGTGTGACCATCAGGTGTGTGATGCTGTGCTGATG 1500
DB 50990 TGTCTGGGAGTGGACAAGAGATGTGTGACCATCAGGTGTGTGATGCTGTGCTGATG 50931
```


Db	48771	AGGGTGATCTCTCGCGCTATGGCCCGCGTGGCGGAGAGAGGCGCTTCTCCGTGTC	48712
Qy	3721	CACCTTGGCAACTTGGGCTGGGCAAGAAGTCGCTGGAGCAGTGGGTGACCGAGGAGGC	3780
Db	48711	CACCTTGGCAACTTGGGCTGGGCAAGAAGTCGCTGGAGCAGTGGGTGACCGAGGAGGC	48652
Qy	3781	CGCTGCTTTTGTCCG- CTTTGGCGNACAAAGCCGGTGGGTGATGGGCAAGAGGCAAA	3839
Db	48651	CGCTGCTTTTGTCCGCGCTTCCGCGACCAAGCCGGTGGGTGATGGGCAAGAGGCAAA	48592
Qy	3840	AGCGGGAATCTGGGAAGCGGGGACGGAGAAGGCAACCCCTTACCCGCACTCTCCCAACC	3899
Db	48591	AGCGGGAATCTGGGAAGCGGGGACGGAGAAGGCAACCCCTTACCCGCACTCTCCCAACC	48533
Qy	3900	CCAGGAGCGCCCTTTTGGCCCAACCGGCTCTTGGACAAAGCCGCTGAGCAACGTGATCGCC	3959
Db	48532	CCAGGAGCGCCCTTTTGGCCCAACCGGCTCTTGGACAAAGCCGCTGAGCAACGTGATCGCC	48473
Qy	3960	TCCCTCACCTTGGGGGCGCGCTTCCAGTACGACGACCCCTGCTTCTCAGGCTGCTGGAC	4019
Db	48472	TCCCTCACCTTGGGGGCGCGCTTCCAGTACGACGACCCCTGCTTCTCAGGCTGCTGGAC	48413
Qy	4020	CTAGCTCAGAGGAGCTGAAGGAGGAGTGGGCTTTCTGCGCGAGGTGCGGAGGAGAGA	4079
Db	48412	CTAGCTCAGAGGAGCTGAAGGAGGAGTGGGCTTTCTGCGCGAGGTGCGGAGGAGAGA	48353
Qy	4080	CCGAGGAGTCTCTGACGGGCGAGCTCCTGAGAGGTGCGGGGCTGGACTGGGGCTCCGA	4139
Db	48352	CCGAGGAGTCTCTGACGGGCGAGCTCCTGAGAGGTGCGGGGCTGGACTGGGGCTCCGA	48293
Qy	4140	AGGGCAGGATTTGCATAGATGGGTTTGGGAAAGGACATTCACGAGACCCCACTGTAAGA	4199
Db	48292	AGGGCAGGATTTGCATAGATGGGTTTGGGAAAGGACATTCACGAGACCCCACTGTAAGA	48233
Qy	4200	AGGGCTTGAGGAGGAGGAGACATCTCAGACATGCTCGTGGGAGAGTGTGCCGGGTCA	4259
Db	48232	AGGGCTTGAGGAGGAGGAGACATCTCAGACATGCTCGTGGGAGAGTGTGCCGGGTCA	48173
Qy	4260	GGGGGACACGAGGAGGAGGAGGAGCTGTGACCCCGCTCCACGTTGAGATTTTCGATTTT	4319
Db	48172	GGGGGACACGAGGAGGAGGAGGAGCTGTGACCCCGCTCCACGTTGAGATTTTCGATTTT	48113
Qy	4320	AGGTTTCTCTCTGGGCAAG- -GAGAGAGGTGGAGGCTGGCACTTTGGGAGGAGCTTG	4377
Db	48112	AGGTTTCTCTCTGGGCAAGAGAGAGAGGCTGGAGGCTGGCACTTTGGGAGGAGCTTG	48053
Qy	4378	TGAGGTCAGTGTGTAAGGACAGGCGCCCTGGGTCTACCTGGAGATGGCTGGGGCTGAG	4437
Db	48052	TGAGGTCAGTGTGTAAGGACAGGCGCCCTGGGTCTTCTGGAGATGGCTGGGGCTGAG	47993
Qy	4438	ACTTGTCCAGGTGAAACGACAGACACAGGAGGATTCAGACCCCGTTCTGTCTGTGTAG	4497
Db	47992	ACTTGTCCAGGTGAAACGACAGACACAGGAGGATTCAGACCCCGTTCTGTCTGTGTAG	47933
Qy	4498	TGCTGAATCTGTCCTCCGCTCTCTCCACATCCAGCGTGGCTGGCAAGGCTCTACGCT	4557
Db	47932	TGCTGAATCTGTCCTCCGCTCTCTCCGCAATCCAGCGTGGCTGGCAAGGCTCTACGCT	47873
Qy	4558	TCCAAAGGCTTCTGACCCAGCTGGATGAGTGTCTAAGTACGACAGGATGACCTGGG	4617
Db	47872	TCCAAAGGCTTCTGACCCAGCTGGATGAGTGTCTAAGTACGACAGGATGACCTGGG	47813
Qy	4618	ACCCAGCCAGCCACCCGAGACCTGACTGAGGCTTCTTGGCAAGAGAGGAGGAGTGA	4677
Db	47812	ACCCAGCCAGCCACCCGAGACCTGACTGAGGCTTCTTGGCAAGAGAGGAGGAGTGA	47753
Qy	4678	GAGTGGCTGCCAGGTGGGGGCAAGGGTGGTGGTTGAACGCTCCAGGAGGAATGAGGG	4737
Db	47752	GAGTGGCTGCCAGGTGGGGGCAAGGGTGGTGGTTGAACGCTCCAGGAGGAATGAGGG	47693
Qy	4738	GAGGCTGGGCAAGGTTGGAACAGTGCATCACCCGCGAGCGGCACTTGGGCTGACAGG	4797
Db	47692	GAGGCTGGGCAAGGTTGGAACAGTGCATCACCCGCGAGCGGCACTTGGGCTGACAGG	47633
Qy	4798	TGCAGAAATTTGAGGTCAATTTGGGGGCTACCCCGTTCTATCCCTGAGTATCTCTCGGCC	4857
Db	47632	TGCAGAAATTTGAGGTCAATTTGGGGGCTACCCCGTTCTATCCCTGAGTATCTCTCGGCC	47573
Qy	4858	CTGCTCAGGCCAAGGGAGGCTCTGAGAGCAGCTTCAATGATGAGAACTCTGCGCATATGG	4917
Db	47572	CTGCTCAGGCCAAGGGAGGCTCTGAGAGCAGCTTCAATGATGAGAACTCTGCGCATATGG	47513
Qy	4918	TGGGTAACCTGTTCTTTCGCGGATGTTGACCACTTCGACCACTGCGGCTTGGGCGCTCC	4977
Db	47512	TGGGTAACCTGTTCTTTCGCGGATGTTGACCACTTCGACCACTGCGGCTTGGGCGCTCC	47453
Qy	4978	TGCTCATGATCTCTACACCTTGGATGTGACGCTGAGCCCAAGCTGGGGCCCAAGGACGGGAC	5037
Db	47452	TGCTCATGATCTCTACACCTTGGATGTGACGCTGAGCCCAAGCTGGGGCCCAAGGACGGGAC	47393
Qy	5038	TGAGGAGGAAGGTTACAGCTGGGGCCCTTGGGCTTAGCTGGGAACCCGGGGCTTCCA	5097
Db	47392	TGAGGAGGAAGGTTACAGCTGGGGCCCTTGGGCTTAGCTGGGAACCCGGGGCTTCCA	47333
Qy	5098	GCACAGCGTGGCCAGGCTCTGTAAGCCTTAACCTTCTCCACACAGGAGGAGGAGAGT	5157
Db	47332	GCACAGCGTGGCCAGGCTCTGTAAGCCTTAACCTTCTCCACACAGGAGGAGGAGAGT	47273
Qy	5158	GTCCCTTGGGTGTGACCCCAATTTGTTGGGAGCGCATGTCTGTCAGTCCGTGTCACACAGGA	5217
Db	47272	GTCCCTTGGGTGTGACCCCAATTTGTTGGGAGCGCATGTCTGTCAGTCCGTGTCACACAGGA	47213
Qy	5218	GATCGACGAGCTGATAGGGAGGTTGGCGGACACAGAGATGGGTGACAGGCTCACATGCC	5277
Db	47212	GATCGACGAGCTGATAGGGAGGTTGGCGGACACAGAGATGGGTGACAGGCTCACATGCC	47153
Qy	5278	CTACACCACTGCGGTGATTACAGAGTGCAGCGCTTTGGGGACATCATCCCCCTGAGTGT	5337
Db	47152	CTGACCACTGCGGTGATTACAGAGTGCAGCGCTTTGGGGACATCATCCCCCTGAGTGT	47093
Qy	5338	GACCCATATGACATCCCGTGACATCGAAGTACAGGCTTCCGCGATCCCTAAGGTAGGCTT	5397
Db	47092	GACCCATATGACATCCCGTGACATCGAAGTACAGGCTTCCGCGATCCCTAAGGTAGGCTT	47033
Qy	5398	GGGGCCCTCTCACCCAGCTCAGACACAGCACTGGTGTATAGCCCGACGATGGCTACTG	5457
Db	47032	GGGGCCCTCTCACCCAGCTCAGACACAGCACTGGTGTATAGCCCGACGATGGCTACTG	46973
Qy	5458	CCAGTGGGCGGCTCTAGGAACCTTGGCCACTAGTCTCAATGACCAACACTGACTG	5517
Db	46972	CCAGTGGGCGGCTCTAGGAACCTTGGCCACTAGTCTCAATGACCAACACTGACTG	46913
Qy	5518	TCCCCCACTTGGGTGGGGGTCCAGAGTATAGCAGGCTGGCTGTCCATCCAGAGCCCTT	5577
Db	46912	TCCCCCACTTGGGTGGGGGTCCAGAGTATAGCAGGCTGGCTGTCCATCCAGAGCCCTT	46853
Qy	5578	CGTCTAGTGGGG- AGACAAACAGGACCTTGGCAGAAATTTGGAGGACCCAGCGCTGCG	5636
Db	46852	AGTCTAGTGGGGAAGACAAATCAGGACCTTGCAGAAATTTGGAGGACCCAGCGCTGCG	46793
Qy	5637	GGAGGGGGGAGTGTGGGTGCTCTGAGAGGTGTGACTGCGCCCTGTCTGTGGGGTCGGA	5696
Db	46792	GGAGGGGGGAGTGTGGGTGCTCTGAGAGGTGTGACTGCGCCCTGTCTGTGGGGTCGGA	46733
Qy	5697	GAGGGTACTGTGTGAGCTTCTCGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAG	5756
Db	46732	GAGGGTACTGTGTGAGCTTCTCGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAG	46673
Qy	5757	GCAGTGTGTCTCCCGTGTGTTTGGTGGCAGGGGTCCAGCATCTCTAGAGTCCAGTCCC	5816
Db	46672	GCAGTGTGTCTCCCGTGTGTTTGGTGGCAGGGGTCCAGCATCTCTAGAGTCCAGTCCC	46613
Qy	5817	CACTCTACCTTGCATCTCTGCCCGAGGAAACGACACTCATCACCAACCTGTCTCGGTG	5876
Db	46612	CACTCTACCTTGCATCTCTGCCCGAGGAAACGACACTCATCACCAACCTGTCTCGGTG	46553

Qy	5877	CTGAAGGATGAGGCCGTCTCGGAGAGACCCCTTCGCTTCCACCCCGAAACATCTCTCTGGAT	5936
Db	46552	CTGAAGGATGAGGCCGTCTGGAAGAGCCCTTCGCTTCCACCCCGAAACATCTCTCTGGAT	46493
Qy	5937	GCCCAGGGCCACATTGTGAAGCCGGAGGCTTCTGCTCTTCTCAGCAGGTGCTGTGGG	5996
Db	46492	GCCCAGGGCCACATTGTGAAGCCGGAGGCTTCTGCTCTTCTCAGCAGGTGCTGTGGG	46433
Qy	5997	GAGCCCGGCTCCCTGTCCCTTCCGTGTGAGTCTTGACGGGTATCACCCAGAGGACGAGC	6056
Db	46432	GAGCCCGGCTCCCTGTCCCTTCCGTGTGAGTCTTGACGGGTATCACCCAGAGGACGAGC	46373
Qy	6057	TCATGACGCCCTCCCTCCCTCCACAGGCGCGTGCATGCTCTGGGGAGCCCTTGCCCC	6116
Db	46372	TCATGACGCCCTCCCTCCCTCCACAGGCGCGTGCATGCTCTGGGGAGCCCTTGCCCC	46313
Qy	6117	GCATGGAGCTCTTCTCTTCTTCACTCCCTGCTGACACATTCAGCTTCTCGTGTGGCG	6176
Db	46312	GCATGGAGCTCTTCTCTTCTTCACTCCCTGCTGACACATTCAGCTTCTCGTGTGGCG	46253
Qy	6177	CCGGACAGCCCCGGCCCCAGCACCTCTCGTGTCTGAGCTTTCTGTGTAGCCCCATCCCT	6236
Db	46252	CCGGACAGCCCCGGCCCCAGCACCTCTCGTGTCTGAGCTTTCTGTGTAGCCCCATCCCT	46193
Qy	6237	ACGAGCTTTGTGTGTGCCCGCTAGAAATGGGTACTAGTCCCAAGCTGCTCCCTAGC	6296
Db	46192	ACGAGCTTTGTGTGTGCCCGCTAGAAATGGGTACTAGTCCCAAGCTGCTCCCTAGC	46133
Qy	6297	CAGAGGCTCTAATGTACAAATAAGCAATGTGTGTAGTTCCAACTTGGGTCCCTGTCTCAG	6356
Db	46132	CAGAGGCTCTAATGTACAAATAAGCAATGTGTGTAGTTCCAACTTGGGTCCCTGTCTCAG	46073
Qy	6357	CCCTCTGTGGATCATCTCTCTCAGGGCAACCCACCCTGCTCTCATCTCGCTTACCCC	6416
Db	46072	CCCTCTGTGGATCATCTCTCTCAGGGCAACCCACCCTGCTCTCATCTCGCTTACCCC	46013
Qy	6417	ACGCCCTGGCCGCAATTTGAGACGGGTACGTTGAGGCTGAGCAGATGTCAGTTACCTTTGC	6476
Db	46012	ACGCCCTGGCCGCAATTTGAGACGGGTACGTTGAGGCTGAGCAGATGTCAGTTACCTTTGC	45953
Qy	6477	CCATAATCCCATGTCCCCCACTGACCCCAACTGTGACTGTGCCCAGATTGGTGAACAAGACTA	6536
Db	45952	CCATAATCCCATGTCCCCCACTGACCCCAACTGTGACTGTGCCCAGATTGGTGAACAAGACTA	45893
Qy	6537	CATTGTCTCGCATGTGGGAAAGGGGCCAGAAATGGGCTGACTAGAGGTGCTGAGTCAGCCC	6596
Db	45892	CATTGTCTCGCATGTGGGAAAGGGGCCAGAAATGGGCTGACTAGAGGTGCTGAGTCAGCCC	45833
Qy	6597	TGGATGTGTGGAGAGGCGCAGGACTCAGCCTCGAGGCCCATATTTTCAGGSCCTAACTCAGC	6656
Db	45832	TGGATGTGTGGAGAGGCGCAGGACTCAGCCTCGAGGCCCATATTTTCAGGSCCTAACTCAGC	45773
Qy	6657	CCACCCCACTACAGGACAGAGTCTTCCAGCACCATCACAAACAGTCACTTCCCTTCAT	6716
Db	45772	CCACCCCACTACAGGACAGAGTCTTCCAGCACCATCACAAACAGTCACTTCCCTTCAT	45713
Qy	6717	ATATGACACCCCAAAATGGAAGACAATCATGTCCAGGAGCTATATGCCAGGSCCTACCTTC	6776
Db	45712	ATATGACACCCCAAAATGGAAGACAATCATGTCCAGGAGCTATATGCCAGGSCCTACCTTC	45653
Qy	6777	CCAGGGCTCAGTCGGCAGGTGCCAGAACATTCCTCTGGGAAGGCCCCAGGAAACCCAGGA	6836
Db	45652	CCAGGGCTCAGTCGGCAGGTGCCAGAACATTCCTCTGGGAAGGCCCCAGGAAACCCAGGA	45593
Qy	6837	CCGAGCCACCGCCCTCAGCCTGTCACTTGTGTCCAAAATTGGTGGGTTCTTGGTCTCAC	6896
Db	45592	CCGAGCCACCGCCCTCAGCCTGTCACTTGTGTCCAAAATTGGTGGGTTCTTGGTCTCAC	45533
Qy	6897	TGACTTCAAGAAATCAAGCCCGTGACCCCTCAGGTGAGTGTACAGTCTTAAAGATGGTG	6956
Db	45532	TGACTTCAAGAAATCAAGCCCGTGACCCCTCAGGTGAGTGTACAGTCTTAAAGATGGTG	45473
Qy	6957	TGTTCAGAGTTTGTCTCTCTGATGTAAAGACGTGTTTCAGAGTTTCTTCTCTCTGGTGGG	7016

45472	Db		TGTTTCAGAGTTTGTTCCTTCTCATGTTAAGACGTGTTTCAGAGTTTCTTCTCTTGGTGCGG	45413
7017	Qy		TGCGTGGCTCTGCTGCGCTTCAGGAGTGAAGCTGCAGACCTTCACAGTCAGTGTGTACGGCT	7076
45412	Db		TGCGTGGCTCTGCTGCGCTTCAGGAGTGAAGCTGCAGACCTTCACAGTCAGTGTGTACGGCT	45353
7077	Qy		CTTAAAGCCTGCACGTACGGAGTGTTCATTCTTCTCGTGGGTGTGTGGTCTCACTGGCC	7136
45352	Db		CTTAAAGCCTGCACGTACGGAGTGTTCATTCTTCTCGTGGGTGTGTGGTCTCACTGGCC	45293
7137	Qy		TCAGGAGTGAACCTGCGAGTCCTTCAGTGTGTACAACTCATAAAGCAGTGTGAGACCCCAAT	7196
45292	Db		TCAGGAGTGAACCTGCGAGTCCTTCAGTGTGTACAACTCATAAAGCAGTGTGAGACCCCAAT	45233
7197	Qy		GAGGAGCAGCAGCAGCAGACCTTACTGCAACACGAAAGAAATGATGGCAACACAGGTTG	7256
45232	Db		GAGGAGCAGCAGCAGCAGACCTTACTGCAACACGAAAGAAATGATGGCAACACAGGTTG	45173
7257	Qy		CCGCTGCTACTTCAGGACGCTGCTTTTATTCCTTATCTGACCCCAACCCACATCCTGC	7316
45172	Db		CCGCTGCTACTTCAGGACGCTGCTTTTATTCCTTATCTGACCCCAACCCACATCCTGC	45113
7317	Qy		TGATTGGCCCAATTTTACAGACAGTGGATTGGTGCATTACAGAGAGCTGATTGTGCAATT	7376
45112	Db		TGATTGGCCCAATTTTACAGACAGTGGATTGGTGCATTACAGAGAGCTGATTGTGCAATT	45053
7377	Qy		TACAAATCCCTGAGCTAGACACAGAGTACTGATTGGTATATTTACAAACCTTTGAGCTAGAC	7436
45052	Db		TACAAATCCCTGAGCTAGACACAGAGTACTGATTGGTATATTTACAAACCTTTGAGCTAGAC	44993
7437	Qy		ACAGAGTGTGTAATGGTGTATTTACAAATCCCTTAGCTAGACATAAAGTGTGTCCAGTCC	7496
44992	Db		ACAGAGTGTGTAATGGTGTATTTACAAATCCCTTAGCTAGACATAAAGTGTGTCCAGTCC	44933
7497	Qy		CCACTAGATTAGCTAGATAGAGTACAGACAGAGACCTGATTGGTGGCTTTTACAAACCTTG	7556
44932	Db		CCACTAGATTAGCTAGATAGAGTACAGACAGAGACCTGATTGGTGGCTTTTACAAACCTTG	44873
7557	Qy		AGTTTACACACAGGGTCTGACTGGTGTGTTTACAAACCTTCAGCTAGACACAGAGTGGCTG	7616
44872	Db		AGTTTACACACAGGGTCTGACTGGTGTGTTTACAAACCTTCAGCTAGACACAGAGTGGCTG	44813
7617	Qy		ATTGGTGTATTTACAACTCTTTTAGCTAGAAAATAAAGTTTCCCAAGTCCCCACACAGATTA	7676
44812	Db		ATTGGTGTATTTACAACTCTTTTAGCTAGAAAATAAAGTTTCCCAAGTCCCCACACAGATTA	44753
7677	Qy		GCTAGATAGAGTGCTAATTGGTGCATGCAACGAACTCGGAGCTAGACACAGAGTGTGTAATT	7736
44752	Db		GCTAGATAGAGTGCTAATTGGTGCATGCAACGAACTCGGAGCTAGACACAGAGTGTGTAATT	44693
7737	Qy		GGTGCTATATCAATCCTCTGGCTAGACATATAAAGTTTCTCAAGTCCCCACCTGACTCAGG	7796
44692	Db		GGTGCTATATCAATCCTCTGGCTAGACATATAAAGTTTCTCAAGTCCCCACCTGACTCAGG	44633
7797	Qy		AGCCACACAGCTTTCGCTAGTGGATCTTATGCCAGGGCCACAGGACAGAGTGCCTGCTA	7856
44632	Db		AGCCACACAGCTTTCGCTAGTGGATCTTATGCCAGGGCCACAGGACAGAGTGCCTGCTA	44573
7857	Qy		GTCCCAACACGGGCACTGTATCTCTAGCCCTTGGGCAAGTGGACGGGACCGAGTGCCTG	7916
44572	Db		GTCCCAACACGGGCACTGTATCTCTAGCCCTTGGGCAAGTGGACGGGACCGAGTGCCTG	44513
7917	Qy		GGAGCAGTGGGAGGCAACCACTCGGGAGGCTCGGGCCTCGCAGGGAGCCCAACCGTAGGGA	7976
44512	Db		GGAGCAGTGGGAGGCAACCACTCGGGAGGCTCGGGCCTCGCAGGGAGCCCAACCGTAGGGA	44453
7977	Qy		GGCTTGGGCATGGCAGGCTGCAAGTCTTAGCCCTTGGCCCGGGGAGGTGACTGAGGCC	8036
44452	Db		GGCTTGGGCATGGCAGGCTGCAAGTCTTAGCCCTTGGCCCGGGGAGGTGACTGAGGCC	44393
8037	Qy		TGCGCACAATTCAGTGTGGTGAAGCCCGGACAGCCAGCTATCTGGGGGACCCCGGTGCC	8096

QY	5278	CTACACCACTGCGGTGATTCACGAGGTGACGGCTTTTGGGACATCATCCCTCAGTGT	5337
Db	20755	CTACACCACTGCGGTGATTCACGAGGTGACGGCTTTTGGGACATCATCCCTCAGTGT	20696
QY	5338	GACCATATGACATCCCGTGAATCGAAGTACAGGCTTCCGATCCCTTAAGTATAGGCT	5397
Db	20695	GACCATATGACATCCCGTGAATCGAAGTACAGGCTTCCGATCCCTTAAGTATAGGCT	20636
QY	5398	GGCGCCCTCTCACCCGACTCAGCACGACACTGGTGTATAGCCCGCAGCATGGTACTG	5457
Db	20635	GGCGCCCTCTCACCCGACTCAGCATCAGC-CCCGGTGGTATAGCCCGCAGCATGGTACTG	20577
QY	5458	CCAGGTGGCCCACTCTAGGAACCTTGGCCACTAGTCTCAATGCCACCACTGACTG	5517
Db	20576	CCAGGTGGCCCACTCTAGGAACCTTGGCCACTAGTCTCAATGCCACCACTGACTG	20517
QY	5518	TCCCACTTTGGGTGGGGGTCCAGAGTATPAGCAGGGCTGGCTGTCCATCCAGAGCCCC	5577
Db	20516	TCCCGCTTGGATGGGGGTCCAGAGTATPAGCAGGGCTGGCTGTCCATCCAGAGCCCC	20457
QY	5578	CGTCTAGTGGGG-AGACAAACAGGACCTGCCAGAAATGTTGAGAACCCAGCGCTGCAG	5636
Db	20456	CGTCTAGTGGGAAGACAAATCAGGACCTGCCAGAAATGTTGAGAACCCAGCGCTGCAG	20397
QY	5637	GGAGAGGGGGCAGTGTGGGTGCTCTGAGAGGTGTGACTGCSCCTGTGTGGGGTCCGA	5696
Db	20396	GGAGAGGGGGCAGTGTGGGTGCTCTGAGAGGTGTGACTGCSCCTGTGTGGGGTCCGA	20337
QY	5697	GAGGGTACTGTGGAGTCTTCCGGGGCGAGGACTAGTTGACAGATCCAGCTGTGTGCCAG	5756
Db	20336	GAGGGTACTGTGGAGTCTTCCGGGGCGAGGACTAGTTGACAGATCCAGCTGTGTGCCAG	20277
QY	5757	GAGTGTGTGTCCCGGTGTGTTGGTGGCAGGGTCCAGCATCTTAGAGTCCAGTCCC	5816
Db	20276	GAGTGTGTGTCCCGGTGTGTTGGTGGCAGGGTCCAGCATCTTAGAGTCCAGTCCC	20217
QY	5817	CACTCTACCCGTGACTCTCCGCCAGGGAAGACACTCATCAACAACTGTCTCGGTG	5876
Db	20216	CACTCTACCCGTGACTCTCCGCCAGGGAAGACACTCATCAACAACTGTCTCGGTG	20157
QY	5877	CTGAAGGATGAGCCGTCTGGGAGAGCCCTTCCGCTTCCACCCCGAACTTCTCGGAT	5936
Db	20156	CTGAAGGATGAGCCGTCTGGGAGAGCCCTTCCGCTTCCACCCCGAACTTCTCGGAT	20097
QY	5937	GGCCAGGGCCATTTGTGAAGCCGAGGCTTCTGCTTCTCAGAGGTGCTGTGGG	5996
Db	20096	GGCCAGGGCCATTTGTGAAGCCGAGGCTTCTGCTTCTCAGAGGTGCTGTGGG	20037
QY	5997	GAGCCGGCTCCTGTCCCTTCCGTGGAGTCTTGCAGGGGTATCACCCAGGAGCCAGGC	6056
Db	20036	GAGCCGGCTCCTGTCCCTTCCGTGGAGTCTTGCAGGGGTATCACCCAGGAGCCAGGC	19977
QY	6057	TCACTGAACCCCTCCTCCCAACAGGCGCGCTGCATGCTCCGGGAGCCCTTGGGCC	6116
Db	19976	TCACTGAACCCCTCCTCCCAACAGGCGCGCTGCATGCTCCGGGAGCCCTTGGGCC	19917
QY	6117	GATGAGTCTTCTCTTTTCACTCCTCTGCTGAGCACTTCAGTTCCTCGTGGCGG	6176
Db	19916	GATGAGTCTTCTCTTTTCACTCCTCTGCTGAGCACTTCAGTTCCTCGTGGCGG	19857
QY	6177	CCGGACAGCCCGGCCCACTCTCGTGTGCTGAGTCTTCTGGTGACCCCATCCCT	6236
Db	19856	CCGGACAGCCCGGCCCACTCTCGTGTGCTGAGTCTTCTGGTGACCCCATCCCT	19797
QY	6237	ACGAGCTTTGTGCTGTGCCCCCTAGAAATGGGGTACCTAGTCCCGAGCCCTGTCCCTAGC	6296
Db	19796	ACGAGCTTTGTGCTGTGCCCCCTAGAAATGGGGTACCTAGTCCCGAGCCCTGTCCCTAGC	19737
QY	6297	CAGAGCTCTAATGTACAAATGAAGCAATGTGTGATGTTCCAACTTGGGTCCCTGTCTACG	6356
Db	19736	CAGAGCTCTAATGTACAAATGAAGCAATGTGTGATGTTCCAACTTGGGTCCCTGTCTACG	19677
QY	6357	CCCTCGTTGGGATCATCTCTCCTCAGGGCAACCCACCCCTGCTCATCTCTTACCCC	6416

Db	19676	CCCTCGTTGGGATCATCTCTCCTCAGGGCAACCCCACTCCTCATCTCTGCTTACCCC	19617
QY	6417	ACCGCTGGCCGATTTGAGACGGGTACGTTGAGGCTGAGCAGATGTCACTTACCTTGC	6476
Db	19616	ACCGCTGGCCGATTTGAGACGGGTACGTTGAGGCTGAGCAGATGTCACTTACCTTGC	19557
QY	6477	CCATAATCCCATGTCCCCCACTGACCCCACTCTGACTGCCAGATTTGTTGACAAAGACTA	6536
Db	19556	CCATAATCCCATGTCCCCCACTGACCCCACTCTGACTGCCAGATTTGTTGACAAAGACTA	19497
QY	6537	CATTGTCTGCTGCTGAGGAGGGGCGAGAAATGGGCTGACTAGAGTGTCACTCAGCCCC	6596
Db	19496	CATTGTCTGCTGCTGAGGAGGGGCGAGAAATGGGCTGACTAGAGTGTCACTCAGCCCC	19437
QY	6597	TGGATGTGGTGGAGGGGACGACTCAGCTCTGGAGGCCATATTTTTCAGGCCCTAACTCAGC	6656
Db	19436	TGGATGTGGTGGAGGGGACGACTCAGCTCTGGAGGCCATATTTTTCAGGCCCTAACTCAGC	19377
QY	6657	CCACCCCATCATCAGGACAGCAGTCTCTGCCAGCACCATCAACAAGTCACTCCTTTCAT	6716
Db	19376	CCACCCCATCATCAGGACAGCAGTCTCTGCCAGCACCATCAACAAGTCACTCCTTTCAT	19317
QY	6717	ATATGACACCCCAAAATGGAAGACAATCATGTCTCAGGGAGCTATATGCCAGGGTACCTC	6776
Db	19316	ATATGACACCCCAAAATGGAAGACAATCATGTCTCAGGGAGCTATATGCCAGGGTACCTC	19257
QY	6777	CCAGGGTCACTCGGCGAGGTGCCAGAACTTCCCTGGGAAGGCCCCAGGAAAAACCCAGGA	6836
Db	19256	CCAGGGTCACTCGGCGAGGTGCCAGAACTTCCCTGGGAAGGCCCCAGGAAAAACCCAGGA	19197
QY	6837	CCGAGCCACCGCCCTCAGCCTGTCACTTGTGTCCAAATTTGGTGGTCTTGTGCTCAC	6896
Db	19196	CCGAGCCACCGCCCTCAGCCTGTCACTTGTGTCCAAATTTGGTGGTCTTGTGCTCAC	19137
QY	6897	TGACTTCAAGAAATCAAGCCGCTGAGTGTTCACAGTGTTCCTTAAAGATGTGTG	6956
Db	19136	TGACTTCAAGAAATCAAGCCGCTGAGTGTTCACAGTGTTCCTTAAAGATGTGTG	19077
QY	6957	TGTTTCAAGTGTGTTCTTCTGATGTTTAAAGACGTGTTTCAAGTGTTCCTTCTGTGGG	7016
Db	19076	TGTTTCAAGTGTGTTCTTCTGATGTTTAAAGACGTGTTTCAAGTGTTCCTTCTGTGGG	19017
QY	7017	TGCGTGTCTTGTGCTGAGTGTGAGGAGTGTGAGGAGTGTTCACAGTGTGTACGGCT	7076
Db	19016	TGCGTGTCTTGTGCTGAGTGTGAGGAGTGTGAGGAGTGTTCACAGTGTGTACGGCT	18957
QY	7077	CTTAAAGGCTGACCGTACGAGTGTTCATTTCTTCTGTTGGTGTGTGCTCTCACTGGCC	7136
Db	18956	CTTAAAGGCTGACCGTACGAGTGTTCATTTCTTCTGTTGGTGTGTGCTCTCACTGGCC	18897
QY	7137	TCAGAGTGAATTCAGTCTCTTCCAGTGTTCACATCAATAAGGAGTGTGGACCCAAAT	7196
Db	18896	TCAGAGTGAATTCAGTCTCTTCCAGTGTTCACATCAATAAGGAGTGTGGACCCAAAT	18837
QY	7197	GAGGAGCAGCAGCAGCAGACTTACTGCAAAACAGCAAAAGAAATGATGGCAACAGAGTGTG	7256
Db	18836	GAGGAGCAGCAGCAGCAGACTTACTGCAAAACAGCAAAAGAAATGATGGCAACAGAGTGTG	18777
QY	7257	CCGCTGTCTTCTCAGGAGCCCTGCTTTTATTTTCCCTTATCTGACCCCAACCACTTCTGCTC	7316
Db	18776	CCGCTGTCTTCTCAGGAGCCCTGCTTTTATTTTCCCTTATCTGACCCCAACCACTTCTGCTC	18717
QY	7317	TGATTTGGCCCAATTTTACAGACAGTGGATGTGCTCACTTACAGAGCTGTGTGGTCAAT	7376
Db	18716	TGATTTGGCCCAATTTTACAGACAGTGGATGTGCTCACTTACAGAGCTGTGTGGTCAAT	18657
QY	7377	TACAACTCCCTGAGCTAGACACAGAGTACTGATTTGGTATATTTTACAACTTGTAGCTAGAC	7436
Db	18656	TACAACTCCCTGAGCTAGACACAGAGTACTGATTTGGTATATTTTACAACTTGTAGCTAGAC	18597
QY	7437	ACAGAGTCTCAATTTGGTGTATTTTACAACTTCTAGCTAGACATAAAGTGTGTCCAGTCC	7496

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1	Heim,M.H. and Meyer,U.A.	Evolution of a highly polymorphic human cytochrome P450 gene	Genomics 14 (1), 49-58 (1992)	93052308	1358797	2 (bases 1 to 13677)	Heim,M.H.	Direct Submission	Submitted (25-MAR-1991) M.H. Heim, Dept of Pharmacology, Biocentre University of Basel, Klingelbergstr 70, 4056 Basel, SWITZERLAND	See X58467, and Am. J. Hum. Genet. 47:994-1001(1990).	Location/Qualifiers	1. .13677	/organism="Homo sapiens"	/mol_type="genomic DNA"	/db_xref="taxon:9606"	/chromosome="22"	/clone="45"	/clone_lib="EMBL4"	/dev_stage="adult"	1534. .5868	/gene="CYP2D7BP"	/pseudo	join(1534. .1801,2504. .2675,3203. .3355,3444. .3605,4031. .4207,4400. .4540,4735. .4922,5377. .5518,5617. .5868)	/gene="CYP2D7BP"	/pseudo	1534. .1801	/gene="CYP2D7BP"	/number=1	/pseudo	join(1622. .1801,2504. .2675,3203. .3355,3444. .3605,4031. .4207,4400. .4540,4735. .4922,5377. .5518,5617. .5795)	/gene="CYP2D7BP"	/pseudo	/codon_start=1	1802. .2503	/gene="CYP2D7BP"	/number=1	2504. .2675	/gene="CYP2D7BP"	/number=2	/pseudo	2676. .3202	/gene="CYP2D7BP"	/number=2	3203. .3355	/gene="CYP2D7BP"	/number=3	/pseudo	3356. .3443	/gene="CYP2D7BP"	/number=3	3444. .3605	/gene="CYP2D7BP"	/number=4	/pseudo	3606. .4030	/gene="CYP2D7BP"	/number=4	4031. .4207	/gene="CYP2D7BP"	/number=5	/pseudo	4208. .4399	/gene="CYP2D7BP"	/number=5	4400. .4540	/gene="CYP2D7BP"	/number=6																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	source	gene	mRNA	exon	CDS	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	intron	exon	in

[illegible]

Qy	2347	ATCAGAGCCAGGCTGAGAGAGGGGAAGAGGTGTTGGGGGACGTTTCTCGGGGAAGGACATT	2406
Db	1890	ACCAGAAGCCAGGCTGAGAGAGGGGAAGAGGTGTTGGGGGAC-TTCTCTGGAAGAAGGGCATT	1948
Qy	2407	TATACATGGCATGAAGACCTGGATTTTCCAAAGGCCAAGGAAGTAGTAGGCAAGGGCCTG	2466
Db	1949	TATACATGGCATGAAGACCTGGATTTTCCAAAGGCCAAGGAAGTAGTAGGCAAGGGCCTG	2008
Qy	2467	GAGGTGGAGCTGGACTTTGGCAGTGGGCATGCAAGCCCATTTGGGCAACATATGTTATGGAG	2526
Db	2009	GAGGTGGAGCTGGACTTTGGCAGTGGGCATGCAAGCCCATTTGGGCAACATATGTTATGGAG	2068
Qy	2527	TACAAAGTCCCTTCTGCTGCACACAGAAAGAAAGCCCTTGGGAATGGAAGATGAGTTAGT	2586
Db	2069	TACAAAGTCCCTTCTGCTGCACACAGAAAGAAAGCCCTTGGGAATGGAAGATGAGTTAGT	2128
Qy	2587	CCTCAGTGGCGTTTAAATACAGAAATCGAGGATGAAGGGGTGCAGTGACCCGGTTCAAA	2646
Db	2129	CCTCAGTGGCGTTTAAATACAGAAATCGAGGATGAAGGGGTGCAGTGACCCGGTTCAAA	2188
Qy	2647	CTTTTGGCACTGTGGGTCTCTGGGCCCTCACTGCTCACCGGCATGGACCATCATCTGGGAA	2706
Db	2189	CTTTTGGCACTGTGGGTCTCTGGGCCCTCACTGCTCACCGGCATGGACCATCATCTGGGAA	2248
Qy	2707	TGGATGCTTAATCGGGGCTCTCGGCAATTTTGGTGACTCTTGGCAAGGTCAATCTGGGT	2766
Db	2249	TGGATGCTTAATCGGGGCTCTCGGCAATTTTGGTGACTCTTGGCAAGGTCAATCTGGGT	2308
Qy	2767	GACGCATCCAAACTGAGTTCTCCATCACAGAGGTGTGACCCCAACCCCTGCCACGA	2826
Db	2309	GACGCATCCAAACTGAGTTCTCCATCACAGAGGTGTGACCCCAACCCCTGCCACGA	2368
Qy	2827	TCAGGAGCTGGGTCTCTCTTCCACCTGTCTCACTCTGTAGTACCCCGGGGTCTGCTCA	2886
Db	2369	TCAGGAGCTGGGTCTCTCTTCCACCTGTCTCACTCTGTAGTACCCCGGGGTCTGCTCA	2428
Qy	2887	AGTTTCAAAATAGACTAGGACTGTAGTCTGGGGTGATCTGTGCTTGACAAGAGGCCCTG	2946
Db	2429	AGTTTCAAAATAGACTAGGACTGTAGTCTGGGGTGATCTGTGCTTGACAAGAGGCCCTG	2488
Qy	2947	ACCTCTCCTCTGCAGTTTCCGGCGCGCTTTCGGGGACGTGTTACGCTGACGCTGGGCTGG	3006
Db	2489	ACCTCTCCTCTGCAGTTTCCGGCGCGCTTTCGGGGACGTGTTTCAGGCTGACGCTGGGCTGG	2548
Qy	3007	ACGCCGTGTGTGCTCAATGGGCTGGCGGCCCTTGGCGAGGCGATGTGTGACCCGCGGC	3066
Db	2549	ACGCCGTGTGTGCTCAATGGGCTGGCGGCCCTTGGCGAGGCGATGTGTGACCCGCGGC	2608
Qy	3067	GAGGACACGGCCGACCGCCGCTGCGCCCATCTACAGGTCTCTGGGCTTCGGGCGCGCT	3126
Db	2609	GAGGACACGGCCGACCGCCGCTGCGCCCATCTACAGGTCTCTGGGCTTCGGGCGCGCT	2668
Qy	3127	TCCCAAGGACGGCGGTGGGGACAGAGACCGCGTTTCCGTGGGCCCCCGGTGACACAG	3186
Db	2669	TCCCAAGGACGGCGGTGGGGACAGAGACCGCGTTTCCGTGGGCCCCCGGTGACACAG	2728
Qy	3187	TGACCGTAGCCCAAGCAGCGCCGACAGGGCGTGGGGTCTCTGGACGTGAAACAGAGATAA	3246
Db	2729	TGACCGTAGCCCAAGCAGCGCCGACAGGGCGTGGGGTCTCTGGACGTGAAACAGAGATAA	2788
Qy	3247	GGCCAGCGAGTGGGCTGAGGACAGTGGGCCAGGAAACCACTCTCACGGGGAGGTGCGAG	3306
Db	2789	GGCCAGCGAGTGGGCTGAGGACAGTGGGCCAGGAAACCACTCTCACGGGGAGGTGCGAG	2848
Qy	3307	TCTGTGGGCTGGHAGGGGGCGGGCTTACTGCCAGAACCCGCGAAGACCCGTTGGGCGAG	3366
Db	2849	TCTGTGGGCTGGHAGGGGGCGG- GGCGTACTGCCACAGACCCGCGAAGACCCGTTGGGCGAG	2907
Qy	3367	GCTCATGCTCGAAGTGGCGTGGCGGGACCGCGGCTATGCTGCGGGCTCAGTGTGGGC	3426
Db	2908	GCTCATGCTCGAAGTGGCGTGGCGGGACCGCGGCTATGCTGCGGGCTCAGTGTGGGC	2967
Qy	3427	GGGACGGGCGGGATCTTCTTTGAGTGGAAAGTGGTCAGGGTGGGACAGACAGAGTGGG	3486

2968	Db	GGGACGGCGCGGATCTTCTTGAGTGAAAGGTGGTCAGGCTGGCGAGACGAGGTGGG	3027
3487	Qy	GCACAAACCCCGCCACGAGGAGGAGCAATGTGGGTGAGCAAAAGAGTGGCCCTGTGCCCC	3546
3028	Db	GCCAAACCCCGCCACGAGGAGGAGCAATGTGGGTGAGCAAAAGAGTGGCCCTGTGCCCC	3087
3547	Qy	AGCTGGACCGGGCTAGGGAATCGCGGAGACCTTTGTGAGCGCCAGGGTTGGAGTGGGTGG	3606
3088	Db	AGCTGGACCGGGCTAGGGAATCGCGGAGACCTTTGTGAGCGCCAGGGTTGGAGTGGGTGG	3147
3607	Qy	CGAGGGTGGGCGAAGGCCCTTCATGGCAACGCCACAGTGTCCGTCCGCCCCCAGGGGT	3666
3148	Db	CGAGGGT- GGGCAAGGCCCTTCATGGCAACGCCACAGTGTCCGTCCGCCCCAAGGGGT	3206
3667	Qy	GATCTCTCGCGCTATGGGCCCGCTGGCGGAGCAGAGCGCTTCTCCGTGTCCACCTT	3726
3207	Db	GATCTCTCGCGCTATGGGCCCGCTGGCGGAGCAGAGCGCTTCTCCGTGTCCACCTT	3266
3727	Qy	GCACAACTTGGGCTTGGGCAAGTAGCTGTGGAGCAGTGGGTGACCGAGAGGCCCGCTG	3786
3267	Db	GCACAACTTGGGCTTGGGCAAGTAGCTGTGGAGCAGTGGGTGACCGAGAGGCCCGCTG	3326
3787	Qy	CCTTTGTGCG- CTTTGC CGACCAAGCGCGTGGTGTATGGGCAAGAGGGCACAAGCGCGG	3845
3327	Db	CCTTTGTGCGCGCTTTCGCCAACCAAGCGGTGGTGTATGGGCAAGAGGGCACAAGCGCGG	3386
3846	Qy	AAC TGGGAAGCGGGGACGGAAGCAACCCCTTACCCCGCATCTCCCCACCCCCAGGA	3905
3387	Db	AAC TGGGAAGCGGGGACGGAAGCGGACCCCTTACCCCGCATCT- CCGACCCCCAGGA	3445
3906	Qy	CGCCCCCTTTCGCCCCCAACGGCCTTTTGGACAAAGCCGTGAGCAACGTGATCGCCTCCCTC	3965
3446	Db	CGCCCCCTTTCGCCCCCAACGGCCTTTTGGACAAAGCCGTGAGCAACGTGATCGCCTCCCTC	3505
3966	Qy	ACCTGCGGGCGCGCTTTCAGTACGACGACCCCTCGCTTCTCAGGCTGCTTGGACCTTAGCT	4025
3506	Db	ACCTGCGGGCGCGCTTTCAGTACGACGACCCCTCGCTTCTCAGGCTGCTTGGACCTTAGCT	3565
4026	Qy	CA- GGAAGGACTCAAGGAGGAGTCGGGCTTCTGCGGAGGTGCGGAGCGAGAGCCGAG	4084
3566	Db	CAGGAGGAGTCAAGGAGGAGTCGGGCTTCTGCGGAGGTGCGGAGC-----AA	3616
4085	Qy	GAGTCTCTGCAGGGCGAGCTCTCTGAGAGGTGCGGGGCTGGACTGGGGGCTCCGAAGGGC	4144
3617	Db	GGGTCTTTGCAGGCGAGCTCTCTGAGAGGTGCGGGGCTGGACTGGGGGCTCCGAAGGGC	3676
4145	Qy	AGGATTTGCATAGATGGTTTGGGAAAGGACATTCAGAGAGACCCCATGTATAGAGGGC	4204
3677	Db	AGGATTTGCATAGATGGTTTGGGAAAGGACATTCAGAGAGACCCCATGTATAGAGGGC	3736
4205	Qy	CTGAGGAGAGGGGACATCTCAGACATGTTCTGTGGGAGAGGTGTCGCCGGGTCAAGGGG	4264
3737	Db	CTGAGGAGAGGGGACATCTCAGACATGTTCTGTGGGAGAGGTGTCGCCGGGTCAAGGGG	3796
4265	Qy	CACGAGAGAGGCCAAGGACTCTGTATCCCGCTCCAGCTTGGAGATTTCCGATTTTAGTTT	4324
3797	Db	CACGAGAGGCCAAGGACTCTGTATCCCGCTCCAGCTTGGAGATTTCCGATTTTAGTTT	3856
4325	Qy	TCTCCTCTGGGCAAG- GAGAGAGGTGGAGGTGCGACATTTGGGAGGGGACTTTGGTGAGG	4382
3857	Db	TCTCCTCTGGGCAAGGAGAGAGGGGTGGAGGCTGGCACTTGGGGAGGGACTTTGGTGAGG	3916
4383	Qy	TCAGTGTAGGACGAGCGGCCCTGGGTCTACTCGAGATGCGCTGGGCGCTGAGACTTG	4442
3917	Db	TCAGTGTGTAGGACGAGCGGCCCTGGGTCTTCTCGAGATGCGCTGGGCGCTGAGACTTG	3976
4443	Qy	TCCAGGTGAACGACGAGCA CAGGAGGATTTGAGACCCCGTTCTGTCTGGTGTAGGTGCTG	4502
3977	Db	TCCAGTGAACGACGAGCA CAGGAGGATTTGAGACCCCGTTCTGTCTGGTGTAGGTGCTG	4036
4503	Qy	AATGCTGTCCCGTCTCTTGCA CATCCAGCGCTGGCTGGCAAGGTCTCTACGTTCCAA	4562

Db 4037 AATGCTGTCCCGTCTCCCGCACATCCAGCGCTGGCTGGCAAGGTCTCTACGCTTCCAA 4096
Qy 4563 AAGGCTTTCTACCCAGCTGATGAGCTGCTAACTGAGCACAGATGACCTGGGACCCA 4622
Db 4097 AAGGCTTTCTACCCAGCTGATGAGCTGCTAACTGAGCACAGATGACCTGGGACCCA 4156
Qy 4623 GCCAGCACACCCCGAGACTGACTAGGCTTTCTGGCAAGAGAGAGAGGTGAGAGTG 4682
Db 4157 GCCAGCACACCCCGAGACTGACTAGGCTTTCTGGCAAGAGAGAGAGGTGAGAGTG 4216
Qy 4683 GCTGCCACGTTGGGGGCGCAAGGTTGGTTGAACTGCTCCAGGAGGAATGAGGGAGGC 4742
Db 4217 GCTGCCACGTTGGGGGCGCAAGGTTGGTTGAACTGCTCCAGGAGGAATGAGGGAGGC 4276
Qy 4743 TGGGCAAAAGGTTGACACGATGATCAACCGGCGAGCGCATCTGGGCTGACAGTGCAG 4802
Db 4277 TGGGCAAAAGGTTGACACGATGATCAACCGGCGAGCGCATCTGGGCTGACAGTGCAG 4336
Qy 4803 AATTGGAGGTCAATTTGGGGGCTACCCGTTCTATCCCTGAGTATCCTCTCGGCCCTGCT 4862
Db 4337 AATTGGAGGTCAATTTGGGGGCTACCCGTTCTATCCCTGAGTATCCTCTCGGCCCTGCT 4396
Qy 4863 CAGGCCAAGGGAGGCTTGAGAGCAGCTTCAATGATGAGAACTTGGCATAGTGGTGGT 4922
Db 4397 CAGGCCAAGGGAGGCTTGAGAGCAGCTTCAATGATGAGAACTTGGCATAGTGGTGGT 4456
Qy 4923 AACCTGTTCTTGGCGGGATGATGACACCTCGAACCGCTGGCTGGGGCTCTCTGCTC 4982
Db 4457 AACCTGTTCTTGGCGGGATGATGACACCTCGAACCGCTGGCTGGGGCTCTCTGCTC 4515
Qy 4983 ATGATCTACCTGATGTCAGCGTGAGCCAGCTGGGGCCCAAGCAGGAGCTGAGG 5042
Db 4516 ATGATCTACCTGATGTCAGCGTGAGCCAGCTGGGGCCCAAGCAGGAGCTGAGG 4575
Qy 5043 GAGGAAGGTTACAGCTGGGGGCTTGGGCTTGGGAGACACCGGGGCTTCCAGACACA 5102
Db 4576 GAGGAAGGTTACAGCTGGGGGCTTGGGCTTGGGAGACACCGGGGCTTCCAGACACA 4635
Qy 5103 GCGTGGCCAGGCTCTGTAAGCTTAACTTCTCTCAACACAGGAGGAAGGAGTGTCC 5162
Db 4636 GCGTGGCCAGGCTCTGTAAGCTTAACTTCTCTCAACACAGGAGGAAGGAGTGTCC 4695
Qy 5163 CTGGGTGTACCCATTTGGGGGCGCATGCTGTCTCAGTCCGTCTCAACAGGAGATCG 5222
Db 4696 CTGGGTGTACCCATTTGGGGGCGCATGCTGTCTCAGTCCGTCTCAACAGGAGATCG 4755
Qy 5223 ACGAGTGTATAGGCGAGTGGCGGACAGAGATGGGTGACAGGCTCACATGCCCTACA 5282
Db 4756 ACGAGTGTATAGGCGAGTGGCGGACAGAGATGGGTGACAGGCTCACATGCCCTACA 4815
Qy 5283 CCACTGCCGTGATTCAGAGGTGCAAGGCTTTGGGGACATCATATCCCTGAGTGTACCC 5342
Db 4816 CCACTGCCGTGATTCAGAGGTGCAAGGCTTTGGGGACATCATATCCCTGAGTGTACCC 4875
Qy 5343 ATATGACATCCCGTACAGTACAGGCTTCCGATCCCTAAGGTAGGCTGGCGC 5402
Db 4876 ATATGACATCCCGTACAGTACAGGCTTCCGATCCCTAAGGTAGGCTGGCGC 4935
Qy 5403 CCTCTCACCCAGCTCAGCACAGCAGTGTGTATAGCCCGAGATGGCTACTGCCAGG 5462
Db 4936 CCTCTCACCCAGCTCAGCACAGCAGTGTGTATAGCCCGAGATGGCTACTGCCAGG 4995
Qy 5463 TGGGCCCACTTAGAAACCTTGGCCACCTTAGTCTCAATGCAACCACTGTATCTGCC 5522
Db 4996 TGGGCCCACTTAGAAACCTTGGCCACCTTAGTCTCAATGCAACCACTGTATCTGCC 5055
Qy 5523 ACTTGGTGGGGGTCAGAGTATAGGAGGCTGGCTGTCTCATCCAGAGCCCGCTCT 5582
Db 5056 ACTTGGTGGGGGTCAGAGTATAGGAGGCTGGCTGTCTCATCCAGAGCCCGCTCT 5115
Qy 5583 AGTGGGGAGACAAACAGGACCTGGCAGAACTTGGAGAGCCAGCGCTGAGGAGAG 5642
Db 5116 AGTGGGGAGACAAACAGGACCTGGCAGAACTTGGAGAGCCAGCGCTGAGGAGAG 5175

Qy 5643 GGGCAGTGTGGGTGCTCTGAGAGTCTGACTGCGCTCTGCTGGGGTTCGGAGAGGCT 5702
Db 5176 GGGCAGTGTGGGTGCTCTGAGAGTCTGACTGCGCTCTGCTGGGGTTCGGAGAGGCT 5235
Qy 5703 ACTGTGAGCTTTCTCGGCGCAGGACTAGTTCACAGAGTCCAGCTGTGTGCGCAGCAGTG 5762
Db 5236 ACTGTGAGCTTTCTCGGCGCAGGACTAGTTCACAGAGTCCAGCTGTGTGCGCAGCAGTG 5295
Qy 5763 TGTGTCCCGCTGTGTGTGTGGTGGCAGGGTCCAGCATCTTAGAGTCCAGTCCCCACTCT 5822
Db 5296 TGTGTCCCGCTGTGTGTGTGGTGGCAGGGTCCAGCATCTTAGAGTCCAGTCCCCACTCT 5355
Qy 5823 CACCTGATCTCTGCGCCAGGACACATCATCACCAACCTGTCTATCGGTGCTGAG 5882
Db 5356 CACCTGATCTCTGCGCCAGGAAACGACATCATCAACCTGTCTATCGGTGCTGAG 5415
Qy 5883 GATGAGGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCCGAACACTTCTTGATGCCAG 5942
Db 5416 GATGAGGCGCTCTGGGAGAGCCCTTCCGCTTCCACCCCGAACACTTCTTGATGCCAG 5475
Qy 5943 GGGCACTTTGTGAAGCCGAGCCCTTCTGCTTTCTCAGCAGTGTCTGTGGGAGGCC 6002
Db 5476 GGGCACTTTGTGAAGCCGAGCCCTTCTGCTTTCTCAGCAGTGTCTGTGGGAGGCC 5535
Qy 6003 GGCTCCCTGTCCCTTCCGTTGAGGCTTTCAGGGGTATCACCCAGGAGCAGGCTCACTG 6062
Db 5536 GGCTCCCTGTCCCTTCCGTTGAGGCTTTCAGGGGTATCACCCAGGAGCAGGCTCACTG 5595
Qy 6063 ACGCCCTTCCCTTCCACACAGCCGCTGATGCTCGGGAGCCCTTGGCCCGCATGG 6122
Db 5596 ACGCCCTTCCCTTCCACACAGCCGCTGATGCTCGGGAGCCCTTGGCCCGCATGG 5655
Qy 6123 AGCTTCTCTTCTTACCTTCCGTTGAGCAGCTTTCAGCTTCTCTGTGGCCGCGGAC 6182
Db 5656 AGCTTCTCTTCTTACCTTCCGTTGAGCAGCTTTCAGCTTCTCTGTGGCCGCGGAC 5715
Qy 6183 AGCCCGGCGCAGCAGCTTCTGCTGCTGCTTCTGCTGAGCCCATCCCTTACGAGC 6242
Db 5716 AGCCCGGCGCAGCAGCTTCTGCTGCTGCTTCTGCTGAGCCCATCCCTTACGAGC 5775
Qy 6243 TTTGTGCTGTGCGCCGCTAGAAATGGGTACCTAGTCCCGCAGCTCTCTAGCCAGAGG 6302
Db 5776 TTTGTGCTGTGCGCCGCTAGAAATGGGTACCTAGTCCCGCAGCTTCTCTAGCCAGAGG 5835
Qy 6303 CTCTAATGTAACAATAAGCAATGTGTAGTTCCTTGGGTCCCTGCTCAAGCCCTCG 6362
Db 5836 CTCTAATGTAACAATAAGCAATGTGTAGTTCCTTGGGTCCCTGCTCAAGCCCTCG 5895
Qy 6363 TTGGATCATCTCTCAGGGCAACCCCAAGCTGCTTATCTGCTTACCCACCCGCT 6422
Db 5896 TTGGATCATCTCTCAGGGCAACCCCAAGCTGCTTATCTGCTTACCCACCCGCT 5955
Qy 6423 TGGCGCATTTTGAAGCGGTAGCTTGGGTGAGCAGATGTACCTTGGTCCCATAA 6482
Db 5956 TGGCGCATTTTGAAGCGGTAGCTTGGGTGAGCAGATGTACCTTGGTCCCATAA 6015
Qy 6483 TCCATGTCTCCCGCACTGACCCCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6542
Db 6016 TCCATGTCTCCCGCACTGACCCCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6075
Qy 6543 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6602
Db 6076 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6135
Qy 6603 TGGTGGAGGCGCAGGACTCAGCTGAGGGCCCATATTTTTCAGGCTTAACTCAGCCACCC 6662
Db 6136 TGGTGGAGGCGCAGGACTCAGCTGAGGGCCCATATTTTTCAGGCTTAACTCAGCCACCC 6195
Qy 6663 CACATCAGGAGCAGAGTCTTCCAGCAGCAGTCAACAGTCACTCTCTCTCATATAGA 6722
Db 6196 CACATCAGGAGCAGAGTCTTCCAGCAGCAGTCAACAGTCACTCTCTCTCATATAGA 6255

QY 6723 CACCCCAAAATGAAGACAAATCATGTGAGGAGCTATATGTCAGGGCTACTCCAGGG 6782
Db 6256 CACCCCAAAATGAAGACAAATCATGTGAGGAGCTATATGTCAGGGCTACTCCAGGG 6315
QY 6783 CTGAGTCGGCAGGTGCCAGAACATTCCTGGGAGGCCCCAGAAAACCCAGACCGAGC 6842
Db 6316 CTGAGTCGGCAGGTGCCAGAACATTCCTGGGAGGCCCCAGAAAACCCAGACCGAGC 6375
QY 6843 CACCGCCCTCAGCCTGTGCTGCTGTCACAAATTCGTGGGTTCTTGGTCTCACTGACTT 6902
Db 6376 CACCGCCCTCAGCCTGTGCTGCTGTCACAAATTCGTGGGTTCTTGGTCTCACTGACTT 6435
QY 6903 CAAGAATGAAGCGTGAGACCTCAGCGTGAGTGTACAGTTCCTAAAGATGTTGTTCA 6962
Db 6436 CAAGAATGAAGCTGTG--ACCTCAGCGTGAGTGTACAGTTCCTAAAGATGTTGTTCA 6493
QY 6963 GAGTTTGTCTTCTGATGTTAAGAGCGTGTTCAGAGTTCTTCTTCTGGTGGTGGTGGT 7022
Db 6494 GAGTTTGTCTTCTGATGTTAAGAGCGTGTTCAGAGTTCTTCTTCTGGTGGTGGTGGT 6553
QY 7023 GTCTTGTGCTTCAGAGTGAAGCTGCAGACCTTCACAGTGAGTGTACGGCTCTTAAG 7082
Db 6554 GTCTTGTGCTTCAGAGTGAAGCTGCAGACCTTCACAGTGAGTGTACGGCTCTTAAG 6613
QY 7083 GCTGCACGTACGGAGTTGTTCAATTCCTTCTGGTGGTGGTGGTGGTGGTGGTGGT 7142
Db 6614 GCTGCACATACGGAGTTGTTCAATTCCTTCTGGTGGTGGTGGTGGTGGTGGTGGT 6673
QY 7143 GTGAACTGAGTTCCTTCAGTGTTCACACTCATAAAGCAGTGTGAGCCCAATGAGGA 7202
Db 6674 GTGAACTGAGTTCCTTCAGTGTTCACACTCATAAAGCAGTGTGAGCCCAATGAGGA 6733
QY 7203 GCAGCAGCAGCAAGCTTACTGCAAAACAGCAAAAGAAATGATGCAACAGGTTGCCGCTG 7262
Db 6734 GCAGCAGCAGCAAGCTTACTGCAAAACAGCAAAAGAAATGATGCAACAGGTTGCCGCTG 6793
QY 7263 CTACTTCAGCAGCCTGCTTTTATTCCTTATCTGACCCCAACCCACATCCTGCTGATTG 7322
Db 6794 CTACTTCAGCAGCCTGCTTTTATTCCTTATCTGACCCCAACCCACATCCTGCTGATTG 6853
QY 7323 GCCATTTTACAGCAGTGTGTTGCTCCACTTACAGAGCTGATTGGTGCAATTTACAAT 7382
Db 6854 GCCATTTTACAGCAGTGTGTTGCTCCACTTACAGAGCTGATTGGTGCAATTTACAAT 6913
QY 7383 CCTGAGCTAGACACAGAGTACTGATTGGTATATTTTACAAACCTTCAGCTAGACACAGAG 7442
Db 6914 CCTGAGCTAGACACAGAGTACTGATTGGTATATTTTACAAACCTTCAGCTAGACACAGAG 6973
QY 7443 TGCTGAATGGTGTATTTACAATCCCTTAGCTAGACATAAAGGTTGTCCAGTCCCCCACTA 7502
Db 6974 TGCTGAATGGTGTATTTACAATCCCTTAGCTAGACATAAAGGTTGTCCAGTCCCCCACTA 7033
QY 7503 GATTAGCTAGATAGTAGACAGAGACACTGATTGGTGGTTTACAACCTTGAATTAG 7562
Db 7034 GATTAGCTAGATAGTAGACAGAGACACTGATTGGTGGTTTACAACCTTGAATTAG 7093
QY 7563 ACACAGGCTGCTGCTGCTGTTTACAACCTTGAAGCTAGACACAGAGTGTGATTGCT 7622
Db 7094 ACACAGGCTGCTGCTGCTGTTTACAACCTTGAAGCTAGACACAGAGTGTGATTGCT 7153
QY 7623 GTATTTACAATCTTTTAGCTAGAAAATAAGGTTTCCCAAGTCCCAACCAAGATTAGCTAGA 7682
Db 7154 GTATTTACAATCTTTTAGCTAGAAAATAAGGTTTCCCAAGTCCCAACCAAGATTAGCTAGA 7213
QY 7683 T--AGAGTGCTAATGGTGCATGCAAGAACCCGAGCTAGACACAGAGTGTGATTGGTG 7740
Db 7214 TACAGAGTGTAAATGGTGCATGCAAGAACCCGAGCTAGACACAGAGTGTGATTGGTG 7273
QY 7741 CATATCAATCTCTGCTAGACATAAAGTTTCCAGTCCCACTGACTCAGAGCC 7800
Db 7274 CATATCAATCTCTGCTAGACATAAAGTTTCCAGTCCCACTGACTCAGAGCC 7333
QY 7801 CAGCCAGCTTCGCTAGTGGATCCTATGTCAGGGCCACAGGCGAGAGCTGCTGTAGTCC 7860

Db 7334 CAGCCAGCTTCGCTAGTGGATCCTATGTCAGGGCCACAGGCGAGCTGCTGTAGTCC 7393
QY 7861 CACACCGGACACTGTGACTCTCAGCCCTTGGGAGTGCAGCGGACACAGTGCCTGGAG 7920
Db 7394 CACACCGGACACTGTGACTCTCAGCCCTTGGGAGTGCAGCGGACACAGTGCCTGGAG 7453
QY 7921 CAGTGGAGGACCCCATCCGCGAGGCTCGGGGCTCGACAGGAGCCACCATGAGGAGGCT 7980
Db 7454 CAGTGGAGGACCCCATCCGCGAGGCTTGGGCTCGACAGGAGCCACCATGAGGAGGCT 7513
QY 7981 TGGGANTGCGAGGCTGCAAGTCTGAGCCCTGCGCCCGGGGAGGTGACTGAGGCTGGC 8040
Db 7514 TGGGANTGCGAGGCTGCAAGTCTGAGCCCTGCGCCCGGGGAGGTGACTGAGGCTGGC 7573
QY 8041 GACAATTCAGGTGTGTGAGCCCGCAGCGCCAGCAGTACTGGGGAGCCCGGTGCCCT 8100
Db 7574 GACAATTCAGGTGTGTGAGCCCGCAGCGCCAGCAGTACTGGGGAGCCCGGTGCCCT 7633
QY 8101 CTGAGCTGCTGGCCCGAGGTGCTAAGCCCTCACTGCTGGGGCCAGAGGACACAGCCGG 8160
Db 7634 CTGAGCTGCTGGCCCGAGGTGCTAAGCCCTCACTGCTGGGGCCAGAGGACACAGCCGG 7693
QY 8161 CCGCTCCGAGTGCAGGCGCCGTGAGCCCTGCGCCCAACCCAGAACTGGTGTGCTGCCCGCA 8220
Db 7694 CCGCTCCGAGTGCAGGCGCCGTGAGCCCTGCGCCCAACCCAGAACTGGTGTGCTGCCCGCA 7753
QY 8221 GCNAACCCAGTTCCCGCACACGCTTCCCTCCATACCTCCCGCAAGCAGACGAGCCG 8280
Db 7754 GCNAACCCAGTTCCCGCACACGCTTCCCTCCATACCTCCCGCAAGCAGACGAGCCG 7813
QY 8281 GCTCCAGCTCCACAGTCCAGAGGCGGCTCCACAGTGCAGCGCTGGGCTGAA 8335
Db 7814 GCTCCAGCTCCACAGTCCAGAGGCGGCTCCACAGTGCAGCGCTGGGCTGAA 7868

RESULT 4
HSCYP2D7A

LOCUS 13278 bp DNA linear PRI 21-OCT-1992
Human CYP2D7A pseudogene for cytochrome P450 2D6.

DEFINITION
X58467.1

Accession X58467.1 GI:30336
Version CYP2D7A gene; Cytochrome P450; cytochrome P450 2D6; pseudogene.

KEYWORDS
SOURCE

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ORGANISM
REFERENCE

1
Heim, M.H. and Meyer, U.A.
Evolution of a highly polymorphic human cytochrome P450 gene

AUTHORS
TITLE

Cluster: CYP2D6
Genomics 14 (1), 49-58 (1992)

JOURNAL
MEDLINE

93052308
PUBMED 1358797

REFERENCE
AUTHORS

1
Heim, M.H.
Direct Submission

TITLE
JOURNAL

Submitted (25-MAR-1991) M.H. Heim, Dept of Pharmacology, Biocentre
University of Basel, Klingelbergstr 70, 4056 Basel, SWITZERLAND
See X58468, and Am. J. Hum. Genet. 47:994-1001(1990).

COMMENT
FEATURES

Location/Qualifiers
1..13278
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="22"
/clone_lib="EMBL4"
/dev_stage="adult"
1154..5489
/gene="CYP2D7A"
/pseudo
Join(1154..1424,2125..2296,2823..2975,3064..3225,

```

3651. .3827,4020. .4161,4356. .4542,4998. .5139,5238. .5489)
/gene="CYP2D7AP"
/pseudo
1154. .1424
/gene="CYP2D7AP"
/number=1
/pseudo
join(1242. .1424,2125. .2296,2823. .2975,3064. .3225,
3651. .3827,4020. .4161,4356. .4542,4998. .5139,5238. .5416)
/gene="CYP2D7AP"
/pseudo
/codon_start=1
1425. .2124
/gene="CYP2D7AP"
/number=1
2125. .2296
/gene="CYP2D7AP"
/number=2
/pseudo
2297. .2822
/gene="CYP2D7AP"
/number=2
2823. .2975
/gene="CYP2D7AP"
/number=3
/pseudo
2976. .3063
/gene="CYP2D7AP"
/number=3
3064. .3225
/gene="CYP2D7AP"
/number=4
/pseudo
3226. .3650
/gene="CYP2D7AP"
/number=4
3651. .3827
/gene="CYP2D7AP"
/number=5
/pseudo
3828. .4019
/gene="CYP2D7AP"
/number=5
4020. .4161
/gene="CYP2D7AP"
/number=6
/pseudo
4162. .4355
/gene="CYP2D7AP"
/number=6
4356. .4542
/gene="CYP2D7AP"
/number=7
/pseudo
4543. .4997
/gene="CYP2D7AP"
/number=7
4998. .5139
/gene="CYP2D7AP"
/number=8
/pseudo
5140. .5237
/gene="CYP2D7AP"
/number=8
5238. .5489
/gene="CYP2D7AP"
/number=9
/pseudo
7887. .7908
repeat_region
/note="ACCCCTCCCC repeat"

```

Query Match 67.2%; Score 6908; DB 9; Length 13278;

Best Local Similarity 99.0%; Pred. No. 0;
Matches 7089; Conservative 0; Mismatches 45; Indels 24; Gaps 13;

QY	1185	TGGGATCCATGTGACAGCTTTGAGGCTCACCGGAGAGACGCTCTGGACAGAGAGGTCCC	1244
DB	349	TAGGATCCACGTGACAGCTTTGAGGCTCACTGGGAGCAGCCTCTGGACAGAGAGGTCCC	408
QY	1245	ATCCAGGAAACCTCGGCGATCGCTGGGAAGTGGGGTACTTGGTCCCGGCTCTGTATGTGT	1304
DB	409	ATCCAGGAAACCTTTGGGATCGCTGGGAAGTGGGGTACTTGGTCCCGGCTCTGTATGTGT	468
QY	1305	CTGTGACCTGGTGTGTGTGAGAGAGAAATGTGTGCCCTGAGTGTCACTGTGAGTCTGTGTAT	1364
DB	469	GTGTGACCTGGTGTGTGTGAGAGAGAAATGTGTGCCCTGAGTGTCACTGTGAGTCTGTGTAT	528
QY	1365	GTGTGAATATTGTCTTTGTGGGTGATTTTCTGCANTGTAAATCGTTCCTGCAAGTG	1424
DB	529	GTGTGAATATTGTCTTTGTGGGTGATTTTCTGCANTGTAAATCGTTCCTGCAAGTG	588
QY	1425	TGAACAAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGGTGTGTC	1484
DB	589	TGAACAAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGGTGTGTC	648
QY	1485	ATAGCGTCTGTGCATGTCAAGAGTCAAGGTGAAGTGAAGGACCCAGGCCCATGTGCCA	1544
DB	649	ATAGCGTCTGTGCATGTCAAGAGTCAAGGTGAAGTGAAGGACCCAGGCCCATGTGCCA	708
QY	1545	CTCATCATCAGGAGCTCTAAGGCCCTAGGTAAAGTGCAGTGCACAGATAAAGGTGTGAAG	1604
DB	709	CTCATCATCAGGAGCTCTAAGGCCCTAGGTAAAGTGCAGTGCACAGATAAAGGTGTGAAG	768
QY	1605	GTCACCTCTGGAGTGGGAGGTGGGGTAGGAAAGGCAAGGTCAATTTCTGGAGGAGGG	1664
DB	769	GTCACCTCTGGAGTGGGAGGTGGGGTAGGAAAGGCAAGGTCAATTTCTGGAGGAGGG	828
QY	1665	GTGTGACTACATAGGGTGTATGAGCCTAGTGGAGGTGGATGGCCGGGTCCACTGAG	1724
DB	829	GTGTGACTACATAGGGTGTATGAGCCTAGTGGAGGTGGATGGCCGGGTCCACTGAG	888
QY	1725	ACCTGGTTATCCAGAAAGCCTGTGTGGGCTTTGGGAGCTTGGAGTGGGAGAGGGGTG	1784
DB	889	ACCTGGTTATCCAGAAAGCCTGTGTGGGCTTTGGGAGCTTGGAGTGGGAGAGGGGTG	948
QY	1785	ACTTCTCCGACAGGCTTTCTACCACTACCTGGTGGTAAAGGCTCTGGAGCAGAGCA	1844
DB	949	ACTTCTCCGACAGGCTTTCTACCACTACCTGGTGGTAAAGGCTCTGGAGCAGAGCA	1008
QY	1845	CGGCAAGGACCTCTGGAGCAGCCCATACCTGCCCTGGCTGACTCTGCCTCTGGCAGCA	1904
DB	1009	CGGCAAGGACCTCTGGAGCAGCCCATACCTGCCCTGACTCTGCCTCTGGCAGCA	1068
QY	1905	CAGTCAACACAGCAGGTTTCACTCAGCAGAGGGGAAAGGCCATCATCAGTCCCTTTAT	1964
DB	1069	CAGTCAACACAGCAGGTTTCACTCAGCAGAGGGGAAAGGCCATCATCAGTCCCTTTAT	1128
QY	1965	AAGGGAAGGTCACGCGCTCGGTGTGCCGAGAGTGTCTGCTGCTGCTCTGTGCTGGT	2024
DB	1129	AAGGGAAGGTCACGCGCTCGGTGTGCCGAGAGTGTCTGCTGCTGCTCTGTGCTGGT	1188
QY	2025	GGGTGGGGGTGCCAGGTGTGTCCAGAGGAGCCCACTTGGTAGTGAGGACCATGGGGC	2084
DB	1189	GGGTGGGGGTGCCAGGTGTGTCCAGAGGAGCCCACTTGGTAGTGAGGACCATGGGGC	1248
QY	2085	TAGAAGCACTGGTGCCTCGCCATGATAGTGGCCATCTTCTGCTCTGTGGAACCTGA	2144
DB	1249	TAGAAGCACTGGTGCCTCGCCATGATAGTGGCCATCTTCTGCTCTGTGGAACCTGA	1308
QY	2145	TGCACCGGACCAAGCTGGGCTGCAAGCTACCGGCGAGGTCCCTTGCACCTGCCGGGC	2204
DB	1309	TGCACCGGACCAAGCTGGGCTGCAAGCTACCGGCGAGGTCCCTTGCACCTGCCGGGC	1368
QY	2205	TGGGCAACCTTGTGTCATGTGGACTTCCAGAAACACCATACTGTTCTGCACAGGTGAGG	2264

ORIGIN

Db 1369 TGGGCAACCTTGCTGATGTGGACTTCCAGAAACACACCATACTGCTTCGACCAAGGTGAGG 1428
QY 2265 GAGGAGGTCTCTGAGGGCGGACAGAGTCTCTGAGGATGCCCAACACAGCAACATATGGT 2324
Db 1429 GAGGAGGTCTCTGAGGGCGGACAGAGTCTCTGAGGATGCCCAACACAGCAACATATGGT 1488
QY 2325 GGTGGGTAAACACACAGGCTGGATCAGAGCCAGGCTGAGAGGGGAAGCAGGTTTGGGG 2384
Db 1489 GGTGGGTAAACACACAGGCTGGATCAGAGCCAGGCTGAGAGGGGAAGCAGGTTTGGGG 1548
QY 2385 GAGCTTCTGGGGAAGGACATTTATACATGGCATGAAGACTGGATTTTCCAAAGGCCAA 2444
Db 1549 GACG-TCTTGGGGAAGACATTTATACATGGCATGAAGACTGGATTTTCCAAAGGCCAA 1607
QY 2445 GGAAGAGTAGGGCAAGGGCTTGAGGTGGAGCTGGACTTGGCAGTGGGATGCAAGCCCA 2504
Db 1608 GGAAGAGTAGGGCAAGGGCTTGAGGTGGAGCTGGACTTGGCAGTGGGATGCAAGCCCA 1667
QY 2505 TTGGGCAACATATGTTATGAGTACAAAGTCCCTTCTGCTGACACACAGAAAGAAAGGCT 2564
Db 1668 TTGGGCAACATATGTTATGAGTACAAAGTCCCTTCTGCTGACACACAGAAAGAAAGGCT 1727
QY 2565 TGGGAATGGAAGATGAGTTAGTCTGAGTGCCGTTTAAATCAGAAATCGAGGATGAAG 2624
Db 1728 TGGGAATGGAAGATGAGTTAGTCTGAGTGCCGTTTAAATCAGAAATCGAGGATGAAG 1787
QY 2625 GGTGACAGTACCCGGTTCAAACCTTTTGACATGTGGTCTCGGGCCTCACTGTCAAC 2684
Db 1788 GGTGACAGTACCCGGTTCAAACCTTTTGACATGTGGTCTCGGGCCTCACTGTCAAC 1847
QY 2685 GGCATGGACCATCATCTGGGATGGGATGCTAACTGGGGCTCTCGGCCAATTTTGGTGAC 2744
Db 1848 GGCATGGACCATCATCTGGGATGGGATGCTAACTGGGGCTCTCGGCCAATTTTGGTGAC 1907
QY 2745 TCTTGAAGGTCACTACTGGGTGACCGCATCAAACTGAGTTCTCCATCAGCAAGAGTGT 2804
Db 1908 TCTTGAAGGTCACTACTGGGTGACCGCATCAAACTGAGTTCTCCATCAGCAAGAGTGT 1967
QY 2805 GACCCACACCTGCCCCAGCATCAGAGGCTGGGTCTCTCTCCATCTGCTCACTCTC 2864
Db 1968 GACCCACACCTGCCCCAGCATCAGAGGCTGGGTCTCTCTCCATCTGCTCACTCTC 2027
QY 2865 TGGTAGCCCCGGGGTCTGTCAGAGTTCAAAATAGGACTAGGACTGTAGTCTGGGGTGT 2924
Db 2028 TGGTAGCCCCGGGGTCTGTCAGAGTTCAAAATAGGACTAGGACTGTAGTCTGGGGTGT 2087
QY 2925 CTTGGCTTGACAGAGGCTCTGACCTCTCTGCAAGTTCGGGCGCGCTTTCGGGACGT 2984
Db 2088 CTTGGCTTGACAGAGGCTCTGACCTCTCTGCAAGTTCGGGCGCGCTTTCGGGACGT 2147
QY 2985 GTTCAGCTGCACTGGGCTGACCGCGGTGCTGCTCAATGGGCTGGGCGCGTGGC 3044
Db 2148 GTTCAGCTGCACTGGGCTGACCGCGGTGCTGCTCAATGGGCTGGGCGCGTGGC 2207
QY 3045 CGAGGCGATGGTGACCGCGGCGGAGACACGGCCGACCGCCGCTGCGCCCATCTACCA 3104
Db 2208 CGAGGCGATGGTGACCGCGGCGGAGACACGGCCGACCGCCGCTGCGCCCATCTACCA 2267
QY 3105 GGTCTTGGGCTTTCGGGCGGTTTCCAAAGCAAGCGCGGTGGGGGACAGACCGGTT 3164
Db 2268 GGTCTTGGGCTTTCGGGCGGTTTCCAAAGCAAGCGCGGTGGGGGACAGACCGGTT 2327
QY 3165 TCCGTGGGCGCGGTGGAGTGAACGATGACCGTACCGCAAGCAGCGCGGAGGGGTG 3224
Db 2328 TCCGTGGGCGCGGTGGAGTGAACGATGACCGTACCGCAAGCAGCGCGGAGGGGTG 2387
QY 3225 CTGGAAGTGAACAGAGATAAGGCCAGCGAGTGGGTGAGGACAGTGGGCCAGGAAAC 3284
Db 2388 CTGGAAGTGAACAGAGATAAGGCCAGCGAGTGGGTGAGGACAGTGGGCCAGGAAAC 2447
QY 3285 ACCTGACCGGGGAGGTGCGAGTCTGTGGGCTGGGAGGGGCGGCTACTGCCAGACC 3344
Db 2448 ACCTGACCGGGGAGGTGCGAGTCTGTGGGCTGGGAGGGGCGG- GGGCTACTGCCAGACC 2506

QY 3345 CGCCAGAAAGCCCGGTGGGCGAGGCTGATCGTTCGAAGTGGCGGTGGCGGGACCGGCGCT 3404
Db 2507 CGCCAGAAAGCCCGGTGGGCGAGGCTGATCGTTCGAAGTGGCGGTGGCGGGG- CGGCGCT 2565
QY 3405 ATGCTGCGGGCTCAGTGTGGGCGGACCGGCGGGATCTTCTTGTAGTGGAAAGGTGGTCA 3464
Db 2566 ATGCTGCGGGCTCAGTGTGGGCGGACCGGCGGGATCTTCTTGTAGTGGAAAGGTGGTCA 2625
QY 3465 GGTGGGCGAGAGACGAGGTGGGCGCAAAACCCGCCCCAGCAGGAGGAGCAATGTGGGTGA 3524
Db 2626 GGTGGGCGAGAGAGGAGTGGGCGCAAAACCCGCCCCAGCAGGAGGAGCAATGTGGGTGA 2685
QY 3525 GCAAGAGTGGGCTCTGTGCCAGCTGGACCGGGCTAGGGACTGCGGGAGACTTGTGGGA 3584
Db 2686 GCAAGAGTGGGCTCTGTGCCAGCTGGACCGGGCTAGGGACTGCGGGAGACTTGTGGGA 2745
QY 3585 GGCACAGGTTGAGTGGGTGGGAGGTTGGGGCCAAAGCCCTTCATGCGCAAGCCCAAG 3644
Db 2746 GGCACAGGTTGAGTGGGTGGGAGGTTGGGGCCAAAGCCCTTCATGCGCAAGCCCAAG 2804
QY 3645 TGTTCGTCCCGCCCGCAGGGGTGATCTGTCGCGCTATGCGGCCCGGCTGCGCGAGCAGA 3704
Db 2805 TGTTCGTCCCGCCCGCAGGGGTGATCTGTCGCGCTATGCGGCCCGGCTGCGCGAGCAGA 2864
QY 3705 GGGCTCTTCTCCGTGTCCACTTTCGCGCAACTTGGGGCTGGGCAAGAGTCCGCTGGAGCAGT 3764
Db 2865 GGGCTCTTCTCCGTGTCCACTTTCGCGCAACTTGGGGCTGGGCAAGAGTCCGCTGGAGCAGT 2924
QY 3765 GGTGACCGAGAGGCGGCTGCTTGTGTCGCG-CTTTCGCGCAAGAGTGGGTGGTGTAT 3823
Db 2925 GGTGACCGAGAGGCGGCTGCTTGTGTCGCGCTTTCGCGCAAGAGTGGGTGGTGTAT 2984
QY 3824 GGGCAGAGGGCACAAAGCGGGAATGGGAAGCGGGGACGAGAGCAACCCCTTAC 3883
Db 2985 GGGCAGAGGGCACAAAGCGGGAATGGGAAGCGGGGACGAGAGCAACCCCTTAC 3044
QY 3884 CGGCATCTCCCAACCCAGGACGCGCCCTTTCGCGCCCAAGCGGCTCTTGGCAAAAGCCGT 3943
Db 3045 CGGCATCT-CCCAACCCAGGACGCGCCCTTTCGCGCCCAAGCGGCTCTTGGCAAAAGCCGT 3103
QY 3944 GAGCAACGTGATCGCTCCCTCAGTGGGCGCGCTTCGAGTACGACGACCTCGCTT 4003
Db 3104 GAGCAACGTGATCGCTCCCTCAGTGGGCGCGCTTCGAGTACGACGACCTCGCTT 3163
QY 4004 CTTAGGCTGCTGGACCTAGCTCA-GGAGGACTGAGAGGAGTGGGCTTTCGCGG 4062
Db 3164 CTTAGGCTGCTGGACCTAGCTCAGGAGGGATTCGAGAGGAGTGGGCTTTCGCGG 3223
QY 4063 AGGTGCGAGCGAGAGACCGAGGAGTCTCTGACGGGCGAGCTCTGAGAGGTGCGGGGC 4122
Db 3224 AGGTGCGAGC-----AAGGCTCTTTCGAGGCGAGCTCTGAGAGGTGCGGGGC 3274
QY 4123 TGGACTGGGCTCTCCGAAGGCGAGGATTTGATAGATGGTTTGGAAAGGACATTCAG 4182
Db 3275 TGGACTGGGCTCTCCGAAGGCGAGGATTTGCTAGATGGTTTGGAAAGGACATTCAG 3334
QY 4183 GAGACCCCACTGTAAAGAGGCTGGAGAGGAGGAGGACATCTCAGACATGTCTGGGA 4242
Db 3335 GAGACCCCACTGTAAAGAGGCTGGAGAGGAGGAGGACATCTCAGACATGTCTGGGA 3394
QY 4243 GAGGTGTGCCGGGTGAGGGGACCCAGGAGAGGCGCAAGGACTCTGTACCCCGTCCACG 4302
Db 3395 GAGGTGTGCCGGGTGAGGGGACCCAGGAGAGGCGCAAGGACTCTGTACCCCGTCCACG 3454
QY 4303 TTGGAGATTTGATTTAGTTTCTCTCTGGGCAAG--GAGAGGGGTGGAGGCTGGCA 4360
Db 3455 TTGGAGATTTGATTTAGTTTCTCTCTGGGCAAGGAGAGGAGGAGGCTGGCA 3514
QY 4361 CTTGGGAGGAGCTTGTGAGGTGAGTAAAGGACAGGCGGCTTGGGTCTACCTGGA 4420
Db 3515 CTTGGGAGGAGCTTGTGAGGTGAGTAAAGGACAGGCGGCTTGGGTCTTCTGGA 3574

QY 4421 GATGGCTGGGGCTGAGACTTGTCCAGGTGAACGACAGACAGAGAGGATTGAGACCC 4480
DB 3575 GATGGCTGGGGCTGAGACTTGTCCAGATGAACGACAGACAGAGAGATTGAGACCCG 3634
QY 4481 GTTCTGTCTGGGTAGTGTGAATGCTGTCTCCCGTCTCTCTGACATCCAGCGCTGGC 4540
DB 3635 GTTCTGTCTGGGTAGTGTGAATGCTGTCTCCCGTCTCTCTGACATCCAGCGCTGGC 3694
QY 4541 TGGCAAGTCTTACGCTTCCAAAGGCTTCTGTGACCCAGCTGGATGAGCTCTAACTGA 4600
DB 3695 TGGCAAGTCTTACGCTTCCAAAGGCTTCTGTGACCCAGCTGGATGAGCTCTAACTGA 3754
QY 4601 GCACAGGATGACTGGGACCCAGCCAGCCAGCCAGCCAGCTGACTGAGCGCTTCTGGC 4660
DB 3755 GCACAGGATGACTGGGACCCAGCCAGCCAGCCAGCCAGCTGACTGAGCGCTTCTGGC 3814
QY 4661 AAAGAAGGAGAGGTGAGTGGCTGCCACGCTGGGGGCAAGGGTGGTGGTGAACGT 4720
DB 3815 AAAGAAGGAGAGGTGAGTGGCTGCCACGCTGGGGGCAAGGGTGGTGGTGAACGT 3874
QY 4721 CCCAGGAGAAATCAGGGGAGGTGGGCAAAAGGTTGGACAGTCAACCCGGGAGCC 4780
DB 3875 CCCAGGAGAAATCAGGGGAGGTGGGCAAAAGGTTGGACAGTCAACCCGGGAGCC 3934
QY 4781 GCATCTGGGCTGACAGGTGCAGAAATGGAGGTCAATTTGGGGGCTACCCCGTTCTATCC 4840
DB 3935 GCATCTGGGCTGACAGGTGCAGAAATGGAGGTCAATTTGGGGGCTACCCCGTTCTATCC 3994
QY 4841 TGAGTATCTCTCGGCCCTCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCAATGATGA 4900
DB 3995 TGAGTATCTCTCGGCCCTCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCAATGATGA 4054
QY 4901 GAACCTCGCATAGTGGGTAACTGTCTTCTCGCGGATGGTGACCACTCGACAC 4960
DB 4055 GAACCTCGCATAGTGGGTAACTGTCTTCTCGCGGATGGTGACCACTCGACAC 4114
QY 4961 GCTGGCTGGGGCTCTGCTCATGATCTACACCTGGATGACGCTGAGCCAGCTG 5020
DB 4115 GCTGGCTGGGGCTCTGCTCATGATCTACACCTGGATGACGCTGAGCCAGCTG 4174
QY 5021 GGGCCCAAGGCAAGGACTGAGGGAGGAGGGTACAGCTGGGGGCCCTGGGCTTAGCTGG 5080
DB 4175 GGGCCCAAGGCAAGGACTGAGGGAGGAGGGTACAGCTGGGGGCCCTGGGCTTAGCTGG 4234
QY 5081 GACACCGGGGCTTCAGACAGGGGTGCGCAGGCTCTGTAAAGCTTAATCTCTCAAC 5140
DB 4235 GACACCGGGGCTTCAGACAGGGGTGCGCAGGCTCTGTAAAGCTTAATCTCTCAAC 4294
QY 5141 ACAGGAGGAGGAGAGTGTCCCTGGGTGCTGACCCATTTGTGGGACGATGCTGTCCA 5200
DB 4295 ACAGGAGGAGGAGAGTGTCCCTGGGTGCTGACCCATTTGTGGGACGATGCTGTCCA 4354
QY 5201 GTCCGTGTCCAAAGGAGATGACGACGCTGATAGGGCAGGTGCGGCGACAGAGATGGGT 5260
DB 4355 GTCCGTGTCCAAAGGAGATGACGACGCTGATAGGGCAGGTGCGGCGACAGAGATGGGT 4414
QY 5261 GACAGGCTCAATGCCCTACACACTGCGCTGATTTACAGAGTGCAGGCTTTGGGAC 5320
DB 4415 GACAGGCTCAATGCCCTACACACTGCGCTGATTTACAGAGTGCAGACTTTGGGAC 4474
QY 5321 ATCATCCCCCTGAGTGTGACCATATGATACCTCCGTGATCGATGACAGGCTTCCGC 5380
DB 4475 -TCGTCCCCCTGGGTGTGACCATATGATACCTCCGTGATCGATGACAGGCTTCCGC 4533
QY 5381 ATCCCCTAAGGTAGGCTGGGCGCCCTCTCACCCAGCTCAGCACAGACACTGGTGATAG 5440
DB 4534 ATCCCCTAAGGTAGGCTGGGCGCCCTCTCACCCAGCTCAGCACAGACCTGGTGATAG 4593
QY 5441 CCCCAGCATGGCTACTGCGAGTGGGCCCACTCTAGGAACTCTAGGCCACCTAGTCTCAA 5500
DB 4594 CCCCAGCATGGCTACTGCGAGTGGGCCCACTCTAGGAACTCTAGGCCACCTAGTCTCAA 4653
QY 5501 TGCCACCACACTGATGTCTCCCACTTTGGTGGGGGCTCCAGAGTATAGGCAGGGCTGGCC 5560

DB 4654 TGCCACCACACTGATGTCTCCCACTTTGGTGGGGGCTCCAGATATAGCAGGGCTGGCC 4713
QY 5561 TGTCCATCCAGAGCCCCGTCTAGTGGGG- AGACAAACAGAGACTCTGCCAGAAATGTGA 5619
DB 4714 TGTCCATCCAGAGCCCCGTCTAGTGGGGAAACAAACAGAGACTCTGCCAGAAATGTGA 4773
QY 5620 GGACCCAGCGCTGCAGGAGAGGGGGCAGTGTGGGTGCTCTGAGAGGTGTGACTGGC 5679
DB 4774 GGACCCAGCGCTGCAGGAGAGGGGGCAGTGTGGGTGCTCTGAGAGGTGTGACTGGC 4833
QY 5680 CCTGTGTGGGTCCGAGAGAGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACAGA 5739
DB 4834 CCTGTGTGGGTCCGAGAGAGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACAGA 4893
QY 5740 GTCCAGCTGTGTGCAGGAGAGTGTGTCTCCCGTGTGTGGTGGGAGGGGTCCAGCA 5799
DB 4894 GTCCAGCTGTGTGCAGGAGAGTGTGTCTCCCGTGTGTGGTGGGAGGGGTCCAGCA 4953
QY 5800 TCCTAGAGTCCAGTCCCACTCTCACCTGATCTCTGCCAGGAAACGACACTCATCA 5859
DB 4954 TCCTAGAGTCCAGTCCCACTCTCACCTGATCTCTGCCAGGAAACGACACTCATCA 5013
QY 5860 CCAACTCTCATCGGTGCTGAAGGATGAGGCGCTCTGGGAGAGGCCCTTCCGCTTCCACC 5919
DB 5014 CCAACTCTCATCGGTGCTGAAGGATGAGGCGCTCTGGAGAGGCCCTTCCGCTTCCACC 5073
QY 5920 CGAACACTTCTCTGGATGCCAGGCGCACTTTGTGAAGCCGAGGCTTCTGCTCTTCT 5979
DB 5074 CGAACACTTCTCTGGATGCCAGGCGCACTTTGTGAAGCCGAGGCTTCTGCTCTTCT 5133
QY 5980 CAGCAGTGTCTGTGGGAGCCGCGCTCTCTGCTCCCTTCCGTGGAGTCTTGCAGGGTA 6039
DB 5134 CAGCAGTGTCTGTGGGAGCCGCGCTCTCTGCTCCCTTCCGTGGAGTCTTGCAGGGTA 5193
QY 6040 TCACCCAGGAGCAGGCTCACTGACGCGCTTCCCTCCCAAGCCGCGGTGATGCT 6099
DB 5194 TCACCCAGGAGCAGGCTCACTGACGCGCTTCCCTCCCAAGCCGCGGTGATGCT 5253
QY 6100 CGGGAGCCCTGGCGCCGATGGAGCTTCTCTTCTTCTACCTCCCTGCTGACGACTT 6159
DB 5254 CGGGAGCCCTGGCGCCGATGGAGCTTCTCTCTTCTTCTACCTCCCTGCTGACGACTT 5313
QY 6160 CAGCTTCTCTGGCGCCGACAGCCCGCGCCAGCCACTCTCTGTCTGCTGAGTCTTCT 6219
DB 5314 CAGCTTCTCTGGCGCCGACAGCCCGCGCCAGCCACTCTCTGTCTGCTGAGTCTTCT 5373
QY 6220 GGTGACCCCATCTCCCTACGAGCTTTGTGCTGTGCTGCTGCTAGAAATGGGGTACCTAGTCC 6279
DB 5374 GGTGACCCCATCTCCCTACGAGCTTTGTGCTGTGCTGCTGCTAGAAATGGGGTACCTAGTCC 5433
QY 6280 CCAGCTGCTCTCTAGCCAGAGGCTCTAATGTACAATAAGCAATGTGGTGTCTCAACT 6339
DB 5434 CCAGCTGCTCTCTAGCCAGAGGCTCTAATGTACAATAAGCAATGTGGTGTCTCAACT 5493
QY 6340 TGGGTCCCTCTCTGACGCGCTCTGTGGGATCATCTCTCTCAGGGGAAACCCACCCCTGCC 6399
DB 5494 CGGGTCCCTCTCTGACGCGCTCTGTGGGATCATCTCTCTCAGGGGAAACCCACCCCTGCC 5553
QY 6400 TCATTCTCTGCTTACCCACCCCTGCGCATTTGAGAGGGGTACGTTGAGGCTCAGCAG 6459
DB 5554 TCATTCTCTGCTTACCCACCCCTGCGCATTTGAGAGGGGTACGTTGAGGCTCAGCAG 5613
QY 6460 ATGTGAGTTACCTTTGGCCATAATCCCAATGCTCCCACTGACCCCACTCTGACTGCCAG 6519
DB 5614 ATGTGAGTTACCTTTGGCCATAATCCCAATGCTCCCACTGACCCCACTCTGACTGCCAG 5673
QY 6520 ATTGTGACAGGACTCAATTGCTCTGCGATGTGGGAGGGGCGGAGAAATGGGCTGACTA 6579
DB 5674 ATTGTGACAGGACTCAATTGCTCTGCGATGTGGGAGGGGCGGAGAAATGGGCTGACTA 5733
QY 6580 GAGGTGTGAGTCAAGCTTGGATGTGGTGGAGGCGAGGACTCAGCTTGGAGGCCCATAT 6639

5734 GAGGTGTAGTCTAGCCCTGGATGTGGTGGAGAGGGCAGGACTCAGCTCAGCTGAGAGCCCATAT 5793
6640 TTCAGGCTTAACCTCAGCCCAACCCACATCAGGAGACAGCAGTCTCGCAGCACCACATCACAA 6699
5794 TTCAGGCTTAACCTCAGCCCAACCCACATCAGGAGACAGCAGTCTCGCAGCACCACATCACAA 5853
6700 CAGTCACCTCCCTTCATATATGACACCCCAAAATGGAAGACAAATCATGTTCAGGAGCTTA 6759
5854 CAGTCACCTCCCTTCATATATGACACCCCAAAATGGAAGACAAATCATGTTCAGGAGCTTA 5913
6760 TATGCCAGGCTACCTCCCAAGGCTCAGTCGGCAGGTGCCAGAACATTCCTCGGAGGC 6819
5914 TATGCCAGGCTACCTCCCAAGGCTCAGTCGGCAGGTGCCAGAACATTCCTCGGAGGC 5973
6820 CCCAGGAAACCCAGGACCGAGCCCTCAGCCCTCAGCTGTGTCACCTTGTGTCACAAATTTG 6879
5974 CCCAGGAAACCCAGGACCGAGCCCTCAGCCCTCAGCTGTGTCACCTTGTGTCACAAATTTG 6033
6880 TGGGTTCTTGGTCTCAGTCTCAAGATGAAGCCGCTGGACCTCAGCTGAGTGTGTTAC 6939
6034 TGGGTTCTTGGTCTCAGTCTCAAGATGAAGCTGTGGACCTCAGCTGAGTGTGTTAC 6093
6940 AGTTCTTAAAGATGCTGTTCAGAGTTTCTTCTGATGTTAAGAGCTGTTCAGAGT 6999
6094 AGTTCTTAAAGATGCTGTTCAGAGTTTCTTCTGATGTTAAGAGCTGTTCAGAGT 6153
7000 TTCTTCTCTTGGTGGTGGTCTGCTGCTTGGCTTCAAGAGTGAAGCTGCAGACCTTCA 7059
6154 TTCTTCTCTTGGTGGTGGTCTGCTGCTTGGCTTCAAGAGTGAAGCTGCAGACCTTCA 6213
7060 CAGTGAAGTTACGGCTTTAAGGCTGCAGTACGAGTGTTCATTTCTTCTGTTGGGT 7119
6214 CAGTGAAGTTACGGCTTTAAGGCTGCAGTACGAGTGTTCATTTCTTCTGTTGGGT 6273
7120 TTGTGCTCTCAGTCTCAGGAGTGAAGCTGAGTCTTCCAGTGTCAACTCATAA 7179
6274 TTGTGCTCTCAGTCTCAGGAGTGAAGCTGAGTCTTCCAGTGTCAACTCATAA 6333
7180 GGCAGTGTGACCCCAATAGGAGGAGCAGCAGCAGAGCTTACTGCAAAACAGCAAAAGAA 7239
6334 GGCAGTGTGACCCCAATAGGAGGAGCAGCAGCAGAGCTTACTGCAAAACAGCAAAAGAA 6393
7240 TGATGGCAACCAAGTTGCCGCTGCTACTTCCAGCAGCTGCTTTTATTTCCCTTATCTGAC 7299
6394 TGATGGCAACCAAGTTGCCGCTGCTACTTCCAGCAGCTGCTTTTATTTCCCTTATCTGAC 6453
7300 CCCCAACCATCTGCTGATTTGGCCCATTTTACAGACAGTGAATGGTCCACTTACAGA 7359
6454 CCCCAACCATCTGCTGATTTGGCCCATTTTACAGACAGTGAATGGTCCACTTACAGA 6513
7360 GAGCTGATTTGCTGATTTACAAATCCCTGAGCTAGACACAGAGTACTGATTTGGTATATTTA 7419
6514 GAGCTGATTTGCTGATTTACAAATCCCTGAGCTAGACACAGAGTACTGATTTGGTATATTTA 6573
7420 CAAACCTTGAGCTAGACACAGAGTGTGATTTGATTTTACAAATCCCTTGTAGTACAT 7479
6574 CAAACCTTGAGCTAGACACAGAGTGTGATTTGATTTTACAAATCCCTTGTAGTACAT 6633
7480 AAGGTTGTCAGTCCCACTAGATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 7539
6634 AAGGTTGTCAGTCCCACTAGATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 6693
7540 TGGGTTTACAAACCTTGAGTTAGACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7599
6694 TGGGTTTACAAACCTTGAGTTAGACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6753
7600 CTAGACACAGAGTGTGATTTGATTTTACAAATCTTTTGTAGCTAGAAATAAGAGTTTCCCT 7659
6754 CTAGACACAGAGTGTGATTTGATTTTACAAATCTTTTGTAGCTAGAAATAAGAGTTTCCCT 6813
7660 AAGTCCCAACAGATAGTGTAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 7717
6814 AAGTCCCAACAGATAGTGTAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 6873

7718 TAGACACAGAGTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 7777
6874 TAGACACAGAGTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 6933
7778 AGTCCCCACCTGACTCAGAGAGCCAGCAGCTTGGCTAGTGGATCTTATGCGAGGCCA 7837
6934 AGTCCCCACCTGACTCAGAGAGCCAGCAGCTTGGCTAGTGGATCTTATGCGAGGCCA 6993
7838 CAGGACAGAGTGTGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 7897
6994 CAGGACAGAGTGTGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 7053
7898 GGAACGAGACAGAGTGTGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 7957
7054 GGAACGAGACAGAGTGTGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 7113
7958 AGGAGACCCACCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 8017
7114 AGGAGACCCACCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 7173
8018 CGGAGAGTGTGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 8077
7174 CGGAGAGTGTGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 7231
8078 TACTTGGGAGACCCGCTGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 8137
7232 TACTTGGGAGACCCGCTGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 7291
8138 CTGGGAGCAGAGACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 8197
7292 CTGGGAGCAGAGACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7351
8198 CCAGAACTGCT 8257
7352 CCAGAACTGCT 7411
8258 CTCCCCGCAAGCAGACCGAGCCGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 8317
7412 CTCCCCGCAAGCAGACCGAGCCGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 7471
8318 GTGACAGCT 8335
7472 GTGACAGCT 7489

RESULT 5
HUMCYP8P
LOCUS
DEFINITION
HUMAN debrisoquine 4-hydroxylase (CYP2D8P) and (CYP2D7) pseudogenes complete sequences.
ACCESSION
M33387
VERSION
M33387.1 GI:181320
KEYWORDS
debrisoquine 4-hydroxylase.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 17060)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Kimura, S., Umeno, M., Skoda, R.C., Meyer, U.A. and Gonzalez, F.J.
AUTHORS
The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and
TITLE
identification of the polymorphic CYP2D6 gene, a related gene, and
a pseudogene
JOURNAL
Am. J. Hum. Genet. 45 (6), 889-904 (1989)
MEDLINE
90072069
PUBMED
2574001
COMMENT
Original source text: Human DNA, clones lambda-2D-A and
lambda-2D-B.
Draft entry and computer-readable sequence for [1] kindly submitted
by S.Kimura, 29-MAR-1990.
Location/Qualifiers
1. .17060
/organism="Homo sapiens"

	/mol_type="genomic DNA"	exon	13147..13307
TATA signal	/db_xref="taxon:9606"		/gene="CYP2D7"
prim_transcript	1276..1282		/number=4
	1304..6570	exon	/pseudo
gene	/note="CYP2D8P mRNA and introns"		13733..13909
	Join(1392..1568,3189..3360,3907..4059,4148..4310,		/gene="CYP2D7"
	4758..4934,5121..5262,5467..5651,6101..6242,6339..6516)		/number=5
CDS	/gene="CYP2D8P"	exon	/pseudo
	Join(1392..1568,3189..3360,3907..4059,4148..4310,		14102..14243
	4758..4934,5121..5262,5467..5651,6101..6242,6339..6516)		/gene="CYP2D7"
	/gene="CYP2D8P"		/number=6
	/codon_start=1	exon	/pseudo
	/product="debrisoquine 4-hydroxylase"		14438..14625
exon	<1392..1568		/gene="CYP2D7"
	/gene="CYP2D8P"		/number=7
exon	/number=1	exon	/pseudo
	3189..3360		15080..15221
	/gene="CYP2D8P"		/gene="CYP2D7"
	/number=2		/number=8
exon	/pseudo	exon	/pseudo
	3907..4059		15320..>15496
	/gene="CYP2D8P"		/gene="CYP2D7"
	/number=3		/number=9
exon	/pseudo		ORIGIN
	4148..4310		Query Match 62.4%; Score 6413.4; DB 9; Length 17060;
	/gene="CYP2D8P"		Best Local Similarity 99.0%; Pred. No. 0;
	/number=4		Matches 6583; Conservative 0; Mismatches 41; Indels 25; Gaps 12;
exon.	/pseudo		
	4758..4934	QY	1185 TGGATCCATGTGACAGCTTTTGAAGCTCACCGGAGACAGCCTCTGGACAGAGAGTGCCC 1244
	/gene="CYP2D8P"	Db	10428 TAGGATCCACGTCACAGCTTTGAGGCTCACCGGAGACAGCCTCTGGACAGAGAGTGCCC 10487
	/number=5		
exon	/pseudo	QY	1245 ATCCAGGAACCTCGGCATGCTGGAGTGGGGTACTTGGTCCGGGTCCTGATGTGT 1304
	5121..5262	Db	10488 ATCCAGGAACCTCGGCATGCTGGAGTGGGGTACTTGGTCCGGGTCCTGATGTGT 10547
	/gene="CYP2D8P"		
	/number=6	QY	1305 GTGTGACTGGTGTGTGAGAGAGAATGTGTGCCCTGAGTGTCAAGTGTGAGTGTCTGTGTAT 1364
exon	/pseudo	Db	10548 GTGTGACTGGTGTGTGAGAGAGAATGTGTGCCCTGAGTGTCAAGTGTGAGTGTCTGTGTAT 10607
	5467..5651	QY	1365 GTGTGAATATTGCTTTTGTGTGGTGTATTTTCGCAATGTGTAATCGTGTCCCTGCAAGTG 1424
	/gene="CYP2D8P"	Db	10608 GTGTGAATATTGCTTTTGTGTGGTGTATTTTCGCAATGTGTAATCGTGTCCCTGCAAGTG 10667
	/number=7		
exon	/pseudo	QY	1425 TGAAACAAGTGGACAAGTGTCTGGAGTGGACAAGAGATCTGTGCACCATCAGTGTGTGC 1484
	6101..6242	Db	10668 TGAAACAAGTGGACAAGTGTCTGGAGTGGACAAGAGATCTGTGCACCATCAGTGTGTGC 10727
	/gene="CYP2D8P"		
	/number=8	QY	1485 ATAGCGTCTGTGTCATGTCGAAGTCAAGTGAAGTGAAGGACCAAGGCCCATGATGCCA 1544
exon	/pseudo	Db	10728 ATAGCGTCTGTGTCATGTCGAAGTCAAGTGAAGTGAAGGACCAAGGCCCATGATGCCA 10787
	6339..6516	QY	1545 CTCATCATCAGAGCTCTAAGCCCCCAGGTAAGTGCAGTGAAGTGCAGATGAAGGGTCTGAAG 1604
TATA signal	/note="CYP2D7 mRNA and introns"	Db	10788 CTCATCATCAGAGCTCTAAGCCCCCAGGTAAGTGCAGTGAAGTGCAGATGAAGGGTCTGAAG 10847
prim_transcript	11236..15571	QY	1605 GTCACCTCTGAGTGGGAGGTGGGGTAGGGAAAGGGCAGGTTCATGTTCTGGAGGAGGG 1664
gene	Join(11324..11504,12206..12377,12906..13058,13147..13307,	Db	10848 GTCACCTCTGAGTGGGAGGTGGGGTAGGGAAAGGGCAGGTTCATGTTCTGGAGGAGGG 10907
	13733..13909,14102..14243,14438..14625,15080..15221,	QY	1665 GTTGTGACTACATTAGGGTGTATGAGCTAGCTGGGAGTGGATGGCCGGGTCCACCTGAG 1724
	15320..15496)	Db	10908 GTTGTGACTACATTAGGGTGTATGAGCTAGCTGGGAGTGGATGGCCGGGTCCACCTGAG 10967
CDS	/gene="CYP2D7"	QY	1725 ACCCTGGTTATCCAGAACCTGTGTGGGCTTGGGAGCTTGGAGTGGGAGAGGGGGTG 1784
	Join(11324..11504,12206..12377,12906..13058,13147..13307,	Db	10968 ACCCTGGTTATCCAGAACCTGTGTGGGCTTGGGAGCTTGGAGTGGGAGAGGGGGTG 11027
	13733..13909,14102..14243,14438..14625,15080..15221,	QY	1785 ACTTCTCCGACGAGGCTTTTCTACCACTACCTACCTGGGTAAAGGCTTGGAGCAGGAGCA 1844
	15320..15496)	Db	11028 ACTTCTCCGACGAGGCTTTTCTACCACTACCTACCTGGGTAAAGGCTTGGAGCAGGAGCA 11087
	/gene="CYP2D7"		
	/number=1		
exon	/product="debrisoquine 4-hydroxylase"		
	<11324..11504		
	/gene="CYP2D7"		
exon	/number=1		
	12206..12377		
	/gene="CYP2D7"		
	/number=2		
exon	/pseudo		
	12906..13058		
	/gene="CYP2D7"		
	/number=3		
	/pseudo		

Qy	1845	GC	GCGCAAGGACCTCTGGAGCAGCCCAATACTGCCCTGGGCTGACTCTG---CCAATGGCA	1901
Db	11088	GC	GCGCAAGGACCTCTGCAGCAGCCCAATACCGGCCCTGGGCTGACCTGCACCCACTGGCA	11147
Qy	1902	GC	CAGTCAACAGCAGGTTCACTCACAGCAGAGGGGAAAGGCCCATCATCAGCTCCCTT	1961
Db	11148	GC	CAGTCAACAGCAGGTTGGCTCACAGCAGAGGGGCAAGGCCCATCATCAGCTCCCTT	11207
Qy	1962	TATA	AGGGAAGGGTCAACGCGCTCGGTGTGCCAGAGTGTCTTGCCTTGGTCTCTCTGTGCT	2021
Db	11208	TATA	AGGGAAGGGTCAACGCGCTCGGTGTGCTGAGAGTGTCTTGCCTCTCTGTGCT	11267
Qy	2022	GGT	GGGTGGGGTGCCAGGTGTGTCCAGAGAGCCCACTGGTAGTGAGGCAGCCATGG	2081
Db	11268	GST	GGGGTGGGGGTGCCAGGTGTGTCCAGAGGAGCCCACTGGTAGTGAGGCAGCCATGG	11327
Qy	2082	GGCT	TAGAAGCACATGTTGCCCTTGGCCCATGATAGTGGCCCATCTTCTGCTCTGTGGACC	2141
Db	11328	GGCT	TAGAAGCACATGTTGGTCCCTTGGCCCATGATAGTGGCCCATCTTCTGCTCTGTGGACC	11387
Qy	2142	TGAT	GCACCGGCACCAACGCTGGGCTGCACGCTTACC CGCAGGTGCCCTGCCCACTGCCCG	2201
Db	11388	TGAT	GCACCGGCACCAACGCTGGGCTGCACGCTTACC CGCAGGTGCCCTGCCCACTGCCCG	11447
Qy	2202	GGCT	TGGGCAACCTTGTGTGCATGTGGACTTCCAGAACACACCATATCTGTTGCAACAGTG	2261
Db	11448	GGCT	TGGGCAACCTTGTGTGCATGTGGACTTCCAGAACACACCATATCTGTTGCAACAGTG	11507
Qy	2262	AGGG	AGAGGTCTTGGAGGGCGCAGAGTCTCTGAGGATGCCCAACCCACAGCAACATG	2321
Db	11508	AGG	AGAGGTCTTGGAGGGCGCAGAGTCTCTGAGGATGCCCAACCCACAGCAACATG	11567
Qy	2322	GTT	GTTGGGTTAAACACACAGGCTTGAATCAGAACGAGCTGAGAAAGGGGAAGCAGTTTG	2381
Db	11568	GTT	GTTGGGTTAAACACACAGGCTTGAATCAGAACGAGCTGAGAAAGGGGAAGCAGTTTG	11627
Qy	2382	GGGA	CGTCTCTGGGAAGGACATTTATATCATGGCATGAAGCATGGATTTTCCAAAGGC	2441
Db	11628	GGG	GAC--TCTTGGGAAGGACATTTATATCATGGCATGAAGCATGGATTTTCCAAAGGC	11685
Qy	2442	CA	AGGAGAGTAGGGCNAGGGCCTGAGGTGGAGCTGGACTTTGGCAGTGGGCATGCAAGC	2501
Db	11686	CA	AGGAGAGTAGGGCAAGGGCCTGAGGTGGAGCTGGACTTTGGCAGTGGGCATGCAAGC	11745
Qy	2502	CCAT	TGGGCAACATATGTTATGAGTACAAAGTCCCTCTCTGTGTGACACCCAGAAAGAAAGG	2561
Db	11746	CCAT	TGGGCAACATATGTTATGAGTACAAAGTCCCTCTCTGTGTGACACCCAGAAAGAAAGG	11805
Qy	2562	CTTT	TGGAAATCGAAGATGAGTTAGTCTCGAGTGCCTTTTAAATACAGAAATCGAGGATGA	2621
Db	11806	CTTT	TGGAAATCGAAGATGAGTTAGTCTCGAGTGCCTTTTAAATACAGAAATCGAGGATGA	11865
Qy	2622	AGG	GGGTGAGTGACCCGGTTTAAACCTTTTGACATGTGGGTCTCTGGGCTCACTGCTC	2681
Db	11866	AGG	GGGTGAGTGACCCGGTTTAAACCTTTTGACATGTGGGTCTCTGGGCTCACTGCTC	11925
Qy	2682	AC	CGGCATGCACCATCATCTGGCAATGGGATGCTAACTGGGGCTCTCGGCAATTTTGGT	2741
Db	11926	AC	CGGCATGCACCATCATCTGGCAATGGGATGCTAACTGGGGCTCTCGGCAATTTTGGT	11985
Qy	2742	GACT	CTTTGCAAGGTTCATCTCTGGGTGACGCATCCAAACTGAGTTCCTCCATCACAGAAAGG	2801
Db	11986	GACT	CTTTGCAAGGTTCATCTCTGGGTGACGCATCCAAACTGAGTTCCTCCATCACAGAAAGG	12045
Qy	2802	TGT	GACCCCAACCGCTGCCCAAGATCAGAGGCTGGGTCTCTCTCTTCCACCTGCTCAC	2861
Db	12046	TGT	GACCCCAACCGCTGCCCAAGATCAGAGGCTGGGTCTCTCTCTTCCACCTGCTCAC	12105
Qy	2862	TCT	TGTAGCCCGGGGGTCTCCAAAGTTTCAATAGGACTAGACCTGTAGTCTGGGGT	2921
Db	12106	TCT	TGTAGCCCGGGGGTCTCCAAAGTTTCAATAGGACTAGGACTGTAGTCTGGGGG	12165
Qy	2922	GAT	CCTGGCTTTGACAAGAGGCCCTGACCTCTCTGTGAGTTTGGCGCGCGCTTCGGGGA	2981

12166	Db		GATCTCTGGCTTGACAAAGAGGCCTTACCCCTCCCTCTGCAGTTGGCGCGCGCTTCGGGGA	12225
2982	Qy	CGTGTTTCAGCTCGAGCTTGGCGCTTGACACGCGTGCTGCTCAATGGGCTTGGCGGCCGT	3041	
12226	Db	CGTGTTTCAGCTCGAGCTTGGCGCTTGACACGCGTGCTGCTCAATGGGCTTGGCGGCCGT	12285	
3042	Qy	GCGCAGGGGATGTGTGAACCCGCGGCGAGGACAACGGCCGACCGCGCTCGCGCCATCTA	3101	
12286	Db	GCSCGAGGCGATGTGTGACCCGCGGCGAGGACAACGGCCGCTCGCGCCATCTA	12345	
3102	Qy	CNAGGTCCTTGGGCTTGGGCGCGGTTCCAAAGGCNAGCGCGGTGGGGGACAGAGACCGC	3161	
12346	Db	CCAAGTCTCTGGGCTTGGGCGCGGTTCCAAAGGCNAGCGCGGTGGGGGACAGAGACCGC	12405	
3162	Qy	GTUUTCCTGGGCGCGGCTGGACAGTGAACGTTAGCCCAAGCAGCGCCGACAGGAGCGCTGGG	3221	
12406	Db	GTUUTCCTGGGCGCGGCTGGACAGTGAACGTTAGCCCAAGCAGCGCCGACAGGAGCGCTGGG	12465	
3222	Qy	GTCTTGAGCGTGAACACAGAGATAAAGGCGACGAGTGGGCTTGAGACAAGTGGGCGCAGAA	3281	
12466	Db	GTCTTGAGCGTGAACACAGAGATAAAGGCGACGAGTGGGCTTGAGACAAGTGGGCGCAGAA	12525	
3282	Qy	ACCACCTGCAACGGGAGGTGCGAGTCTGTGGGCTTGGAGGGGCGGGCTACTTGCCTAG	3341	
12526	Db	ACCACCTGCAACGGGAGGTGCGAGTCTGTGGGCTTGGAGGGGCGGGCTACTTGCCTAG	12585	
3342	Qy	ACCCGCGCAGAAGCCCGGTGGGCGAGGCTGATGCTGCAAGTGGCGGTGGCGGGGACCGCG	3401	
12586	Db	ACCCGCGCAGAAGCCCGGTGGGCGAGGCTGATGCTGCAAGTGGCGGTGGCGGGGACCGCG	12644	
3402	Qy	CCTATCTCTCGGGCTCAGTGTGGGCGGGAACGGGCGGGATCTTCTTGAAGTGGAAAGGTGG	3461	
12645	Db	CCTATCTCTCGGGCTCAGTGTGGGCGGGAACGGGCGGGATCTTCTTGAAGTGGAAAGGTGG	12704	
3462	Qy	TCAGGCTGGGCAAGACGAGGTGGGGCCAAACCCCGCCCAAGCAGGAGGAGCAATGTGGG	3521	
12705	Db	TCAGGCTGGGCAAGACGAGGTGGGGCCAAACCCCGCCCAAGCAGGAGGAGCAATGTGGG	12764	
3522	Qy	TGAGCAAGAGGTGGGCGCTGTGCCAGCTTGAGCCGGGCTAGGGAATCGCGGAGACCTTGT	3581	
12765	Db	TGAGCAAGAGGTGGGCGCTGTGCCAGCTTGAGCCGGGCTAGGGAATCGCGGAGACCTTGT	12824	
3582	Qy	GGAGCCACAGGTTGAGTGGGTGGCGAGGTTGGGGCCAAAGGCTTCATGGCAACGCCC	3641	
12825	Db	GGAGCCACAGGTTGAGTGGGTGGCGAGGTTGGGGCCAAAGGCTTCATGGCAACGCCC	12884	
3642	Qy	ACGTGTCCGTCCCGCCCAAGGGTGATCTCTGTCCGCTATATGGGCCCGCTGGCGCGAGC	3701	
12885	Db	ACGTGTCCGTCCCGCCCAAGGGTGATCTCTGTCCGCTATATGGGCCCGCTGGCGCGAGC	12944	
3702	Qy	AGAGGCGCTTCCGTGTCACCTTTGGCCACTTGGGCTGGGCAAGAGTCSGCTGGAGC	3761	
12945	Db	AGAGGCGCTTCCGTGTCACCTTTGGCCACTTGGGCTGGGCAAGAGTCSGCTGGAGC	13004	
3762	Qy	AGTGGGTGACGAGGAGGCGGCTGCCCTTTGTGCGCG - CTTGCGCGACCAAGCCGCTGGGT	3820	
13005	Db	AGTGGGTGACGAGGAGGCGGCTGCCCTTTGTGCGCGCTTTCGCGCAACCAAGCCGCTGGGT	13064	
3821	Qy	GATGGGCAAGGGCAAAAGCGGAACTTGGGAAGGCGGGGGAACGGAAGAGCCCAACCCCT	3880	
13065	Db	GATGGGCAAGGGCAACAGCGGAACTTGGGAAGGCGGGGGAACGGAAGAGCCCAACCCCT	13124	
3881	Qy	TACCCGCATCTCCCAACCCCAAGGAGCGCCCTTTGCGCCCAACCGGCTCTTTGGACAAGC	3940	
13125	Db	TACCCGCATCT - CCAACCCCAAGGAGCGCCCTTTTTCGCCCAACCGGCTCTTTGGACAAGC	13183	
3941	Qy	CGTGAGCAACGTGATCGCTCCCTCACTGTGGGGCGCGCTTCGAGTACGACGACCCCTCG	4000	
13184	Db	CGTGAGCAACGTGATCGCTCCCTCACTGTGGGGCGCGCTTCGAGTACGACGACCCCTCG	13243	
4001	Qy	CTTCTCTCAGGCTGCTGGAACCTAGCTCAGAGGGGACTGAAGGAGGAGTGGGCTTTCTGCG	4060	

Db	13244	CTTCTCAGGCTGCTGGACCTAGCTCAGGAGGACCTGAAGGAGGAGTCCGGCTTCTCTGCG	13303
Qy	4061	CGAGGTGGGAGGAGAGACCGAGGAGTCTCTGACGGGCGAGCTCTGAGAGGTCCCGGG	4120
Db	13304	CGAGGTGGGAGC-----AAGGGCTTTTGCAGGGCGAGCTCTGAGAGGTCCCGGG	13354
Qy	4121	GCTGACGTGGGCTCCGAAGGCGCAGGATTTGCATAGATGGGTTTGGGAAAGGACATTCC	4180
Db	13355	GCTGACGTGGGCTCCGAAGGCGCAGGATTTGCGTAGATGGGTTTGGGAAAGGACATTCC	13414
Qy	4181	AGGAGACCCACTGTAAAGAGGCTCTGGAGGAGGAGGACATCTCAGACATGGTCTGG	4240
Db	13415	AGGAGACCCACTGTAAAGAGGCTCTGGAGGAGGAGGAGACATCTCAGACATGGTCTGG	13474
Qy	4241	GAGAGGTGTCCCGGTGAGGGGACACAGGAGAGGCCAAGAGCTCTGTACCCCGTCCA	4300
Db	13475	GAGAGGTGTCCCGGTGAGGGGACACAGGAGAGGCCAAGAGCTCTGTACCCCGTCCA	13534
Qy	4301	CGTTGGAGATTTTCGATTTTTCCTCTCTGGGCAAG--GAGAGAGGTTGGAGGCTGG	4358
Db	13535	CGTTGGAGATTTTCGATTTTTCCTCTCTGGGCAAGGAGAGAGAGGTTGGAGGCTGG	13594
Qy	4359	CACTTGGGAGGACCTTGGTAGGTCAAGGTAAAGGACAGGAGGCCCTTGGGTCTACCTG	4418
Db	13595	CACTTGGGAGGACCTTGGTAGGTCAAGGTAAAGGACAGGAGGCCCTTGGGTCTTCTCTG	13654
Qy	4419	GAGATGGCTGGGCTGAGACTTGTCCAGGTGAAACGAGAGACAGGAGGAGTTGAGACC	4478
Db	13655	GAGATGGCTGGGCTGAGACTTGTCCAGGTGAAACGAGAGACAGGAGGAGTTGAGACC	13714
Qy	4479	CCGTTCTGTCTGGTGTAGGTGCTGAATGTGTCTCCCTCTCTGACACATCCACGGCTG	4538
Db	13715	CCGTTCTGTCTGGTGTAGGTGCTGAATGTGTCTCCCTCTCTGACACATCCACGGCTG	13774
Qy	4539	GCTGGCAGGCTCTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGAGTGTAACT	4598
Db	13775	GCTGGCAGGCTCTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGAGTGTAACT	13834
Qy	4599	GAGCAGAGTACCTTGGAGCCAGCCAGCCAGCCAGCTGAGTGGAGGCTTCTCTG	4658
Db	13835	GAGCAGAGTACCTTGGAGCCAGCCAGCCAGCCAGCTGAGTGGAGGCTTCTCTG	13894
Qy	4659	GCAAGAGAGGAGTGAAGTGTGCTGCGAGGTGGGGGCAAGGGTGGTGGTGAAC	4718
Db	13895	GCAAGAGTGAAGTGAAGTGTGCTGCGAGGTGGGGGCAAGGGTGGTGGTGAAC	13954
Qy	4719	GTCCAGGAGGAATGAGGGAGGCTGGGCAAAAGGTTGGACCAAGTGCATCACCCGGCGAG	4778
Db	13955	GTCCAGGAGGAATGAGGGAGGCTGGGCAAAAGGTTGGACCAAGTGCATCACCCGGCGAG	14014
Qy	4779	CCGATCTGGGCTGACAGGTGAGAAATTGGAGGTCAATTTGGGGGTACCCCGTTCTATCC	4838
Db	14015	CCGATCTGGGCTGACAGGTGAGAAATTGGAGGTCAATTTGGGGGTACCCCGTTCTATCC	14074
Qy	4839	CTGAGTATCTCTCGGCCCTCTCAGGCCAAGGGAGCCCTGAGAGAGGCTTCAATGAT	4898
Db	14075	CTGAGTATCTCTCGGCCCTCTCAGGCCAAGGGAGCCCTGAGAGAGGCTTCAATGAT	14134
Qy	4899	GAGAACCTCGCATAGTGGTGGTAACTTCTTCTGGCGGATGGTGACACCTCTGACC	4958
Db	14135	GAGAACCTCGCATAGTGGTGGTAACTTCTTCTGGCGGATGGTGACACCTCTGACC	14194
Qy	4959	AGCTGGCTTGGGGCTCTCTCATGATCTTACACCTTGGATGTGAGCGGTGAGGCCAGC	5018
Db	14195	AGCTGGCTTGGGGCTCTCTCATGATCTTACACCTTGGATGTGAGCGGTGAGGCCAGC	14254
Qy	5019	TGGGGCCCAAGCAGGACTGAGGAGGAGGAGGTACAGTGGGGGCCCTTGGGCTTAGCT	5078
Db	14255	TGGGGCCCAAGCAGGACTGAGGAGGAGGAGGTACAGTGGGGGCCCTTGGGCTTAGCT	14314
Qy	5079	GGGACACCCGGGCTTCCAGCAGCAGCGGTGGCCAGGCTCTCTGAAGCCCTAACTTCTCCA	5138
Db	14315	GGGACACCCGGGCTTCCAGCAGCAGCGGTGGCCAGGCTCTCTGAAGCCCTAACTTCTCCA	14374
Qy	5139	ACACAGGAGGAGGAGAGTGTCCCTGGGTGCTGACCCATTGTGGGGACGATGTCTGTC	5198
Db	14375	ACACAGGAGGAGGAGAGTGTCCCTGGGTGCTGACCCATTGTGGGGACGATGTCTGTC	14434
Qy	5199	CAGTCCGTGTCCAAACAGGAGATCGACACGCTGATAGGGCAGGTGCGGCGACAGAGATGG	5258
Db	14435	CAGTCCGTGTCCAAACAGGAGATCGACACGCTGATAGGGCAGGTGCGGCGACAGAGATGG	14494
Qy	5259	GTGACACAGGCTCAATGCCCCTACACCACTGCGGTGATTTCAAGAGGTGAGCGCTTTGGGG	5318
Db	14495	GTGACACAGGCTCAATGCCCCTACACCACTGCGGTGATTTCAAGAGGTGAGCGCTTTGGGG	14554
Qy	5319	ACATCATCCCCCTGAGTGTGACCCATATGACATCCCGTGACATCGAAGTACAGGCTTCC	5378
Db	14555	ACATCATCCCCCTGAGTGTGACCCATATGACATCCCGTGACATCGAAGTACAGGCTTCC	14614
Qy	5379	GCATCCCTAAGGTAGGCTGGCGCCCTCTCACCCAGCTCAGCACACACCTTGGTGAT	5438
Db	14615	GCATCCCTAAGGTAGGCTGGCGCCCTCTCACCCAGCTCAGCATCAGC-CCCGTGGT	14673
Qy	5439	AGCCCCAGCATGGCTACTGCGCAGGTGGGCCCACTCTAGGAAACCTTGGCCACTAGTCTC	5498
Db	14674	AGCCCCAGCATGGCTACTGCGCAGGTGGGCCCACTCTAGGAAACCTTGGCCACTAGTCTC	14733
Qy	5499	AATGCCACCACTGACCTGCTCCCACTTGGGTGGGGGTCCAGAGTATAGCAGGCGCTGG	5558
Db	14734	AATGCCACCACTGACCTGCTCCCACTTGGGTGGGGGTCCAGAGTATAGCAGGCGCTGG	14793
Qy	5559	CCTGTCCATCAGAGCCGCCCTCTAGTGGGG-AGACAAACACGAGACCTTCCAGAAATGTTG	5617
Db	14794	CCTGTCCATCAGAGCCGCCCTCTAGTGGGGAAACAAATCAGGACCTTCCAGAAATGTTG	14853
Qy	5618	GAGACCCAGCGCTGACAGGAGAGGGGCAAGTGTGGGTGCTCTGAGAGGTGACTGTC	5677
Db	14854	GAGACCCAGCGCTGACAGGAGAGGGGCAAGTGTGGGTGCTCTGAGAGGTGACTGTC	14913
Qy	5678	GCCTGTCTGGGGTCCGAGAGGGTACTGTGAGAGCTTCTCGGGCGCAGGACTAGTTGACA	5737
Db	14914	GCCTGTCTGGGGTCCGAGAGGGTACTGTGAGAGCTTCTCGGGCGCAGGACTAGTTGACA	14973
Qy	5738	GAGTCCAGCTGTGTCAGGCGAGTGTGTCCCCCGTGTGTGTGGTGGAGGGTCCCAG	5797
Db	14974	GAGTCCAGCTGTGTCAGGCGAGTGTGTCCCCCGTGTGTGTGGTGGAGGGTCCCAG	15033
Qy	5798	CATCTAGAGTCCAGTCCCACTCTCACCTTGCATCTCTCCCGCAGGAAACGACACTCAT	5857
Db	15034	CATCTAGAGTCCAGTCCCACTCTCACCTTGCATCTCTCCCGCAGGAAACGACACTCAT	15093
Qy	5858	CACCAAACCTGTCTGCTGAGGATGAGGCGCTCTGGGAGAGCCCTTCCGCTTCCA	5917
Db	15094	CACCAAACCTGTCTGCTGAGGATGAGGCGCTCTGGGAGAGCCCTTCCGCTTCCA	15153
Qy	5918	CCCCGAAACACTTCTGGATGCCCCAGGGGCACTTTGTGAAGCGGAGGCTTCTTCCCTTT	5977
Db	15154	CCCCGAAACACTTCTGGATGCCCCAGGGGCACTTTGTGAAGCGGAGGCTTCTTCCCTTT	15213
Qy	5978	CTCAGCAGGTGCTGTGGGGAGCCCGCTCCCTGTCCCTTCCGTGGAGTCTTGCAGGGG	6037
Db	15214	CTCAGCAGGTGCTGTGGGGAGCCCGCTCCCTGTCCCTTCCGTGGAGTCTTGCAGGGG	15273
Qy	6038	TATCACCCAGGAGCCAGGCTCAGTACGCCCTCTCCCTCCCAAGGCGCGGTGATGC	6097
Db	15274	TATCACCCAGGAGCCAGGCTCAGTACGCCCTCTCCCTCCCAAGGCGCGGTGATGC	15333
Qy	6098	CTCGGGAGCCCTTGGGCCCGCATGAGCTCTTCTCTTTCACCTCCTCTGCTGAGCAC	6157
Db	15334	CTCGGGAGCCCTTGGGCCCGCATGAGCTCTTCTCTTTCACCTCCTCTGCTGAGCAC	15393
Qy	6158	TTCAGCTTCTCCGTGGCCCGGAGAGCCCGGGCCAGCCACTCTCTGCTGCTGAGCTTT	6217
Db	15394	TTCAGCTTCTCCGTGGCCCGGAGAGCCCGGGCCAGCCACTCTCTGCTGCTGAGCTTT	15453

```
QY 6218 CTGGTGACCCCATCCCTTACAGAGCTTTGTGCTGTGCGCCCGCTAGAAATGGGGTACCTAGT 6277
DB 15454 CTGGTGACCCCATCCCTTACAGAGCTTTGTGCTGTGCGCCCGCTAGAAATGGGGTACCTAGT 15513
QY 6278 CCCAGCTGCTCCCTAGCCAGAGGCTCTAAATGTACAATAAAGCAATGTGTAGTTCCAA 6337
DB 15514 CCCAGCTGCTCCCTAGCCAGAGGCTCTAAATGTACAATAAAGCAATGTGTAGTTCCAA 15573
QY 6338 CTTGGGTCCCTGCTCAAGCCTCTGTTGGGATCATCTCTCAGAGGCAACCCACCCCTG 6397
DB 15574 CTTGGGTCCCTGCTCAAGCCTCTGTTGGGATCATCTCTCAGAGGCAACCCACCCCTG 15633
QY 6398 CCTCATCTCTGCTTACCCCAACCCCTGCGCGCATTTGAGAGCGGTAAGTTGAGGCTGAGC 6457
DB 15634 CCTCATCTCTGCTTACCCCAACCCCTGCGCGCATTTGAGAGCGGTAAGTTGAGGCTGAGC 15693
QY 6458 AGATGTCAGTTACCTTTGCCCATTAATCCCATGTCCCCCTGACCTGACCCAACTCTGACCTGCC 6517
DB 15694 AGATGTCAGTTACCTTTGCCCATTAATCCCATGTCCCCCTGACCTGACCCAACTCTGACCTGCC 15753
QY 6518 AGATTGGTGACAAAGGACTACATTTGCTCTGTCATGTGGGAAAGGGGCCAGAAATGGGCTGAC 6577
DB 15754 AGATTGGTGACAAAGGACTACATTTGCTCTGTCATGTGGGAAAGGGGCCAGAAATGGGCTGAC 15813
QY 6578 TAGAGTGTCAAGTCAGCCCTGATGTGTGGAGAGGCGAGGACTCAGCTTGGAGGCCCAT 6637
DB 15814 TAGAGTGTCAAGTCAGCCCTGATGTGTGGAGAGGCGAGGACTCAGCTTGGAGGCCCAT 15873
QY 6638 ATTTAGGCGCTAACTCAGCCCAACCCACATCAGGACAGGAGCTCTGCCAGCACCATCAC 6697
DB 15874 ATTTAGGCGCTAACTCAGCCCAACCCACATCAGGACAGGAGCTCTGCCAGCACCATCAC 15933
QY 6698 AACAGTCACCTCCCTTTCATATATGACACCCCAAATGGAAGACAAATCATGTGAGGAGC 6757
DB 15934 AACAGTCACCTCCCTTTCATATATGACACCCCAAATGGAAGACAAATCATGTGAGGAGC 15993
QY 6758 TATATGCCAGGGCTACTCCAGGGCTCAGTCGGCAGGTGCCAGAACTTCCCTGGGAAG 6817
DB 15994 TATATGCCAGGGCTACTCCAGGGCTCAGTCGGCAGGTGCCAGAACTTCCCTGGGAAG 16053
QY 6818 GCCCAGGAAACCCAGGACCCAGCCACCCCTCAGCTGCTGACCTGTGTGTCAAAAT 6877
DB 16054 GCCCAGGAAACCCAGGACCCAGCCACCCCTCAGCTGCTGACCTGTGTGTCAAAAT 16113
QY 6878 GGTGGGTCTTGTGCTCACTCACTTCAAGATGAAGCCGTGGACCTCAGCGTGAAGTGT 6937
DB 16114 GGTGGGTCTTGTGCTCACTCACTTCAAGATGAAGCCGTGGACCTCAGCGTGAAGTGT 16173
QY 6938 ACAGTTCTTAAAGATGGTGTTCAGAGTTGTCTTCTGATGTTAAGACGTGTTTCA 6997
DB 16174 ACAGTTCTTAAAGATGGTGTTCAGAGTTGTCTTCTGATGTTAAGACGTGTTTCA 16233
QY 6998 GTTCTTCTTCTGTTGGGTGGTGTCTGCTGCTTGGCTTCAAGAGTGAAGCTGAGACCTT 7057
DB 16234 GTTCTTCTTCTGTTGGGTGGTGTCTGCTGCTTGGCTTCAAGAGTGAAGCTGAGACCTT 16293
QY 7058 CACAGTGAAGTTTACGGCTCTTAAAGCTCAGCTCAGGAGTGTTCATCTTCTGTTGG 7117
DB 16294 CACAGTGAAGTTTACGGCTCTTAAAGCTCAGCTCAGGAGTGTTCATCTTCTGTTGG 16353
QY 7118 GTTGTGTGCTCACTGGCCCTCAGGAGTGAATCTGAGTCTTCCAGTGTTTACAACCTCAT 7177
DB 16354 GTTGTGTGCTCACTGGCCCTCAGGAGTGAATCTGAGTCTTCCAGTGTTTACAACCTCAT 16413
QY 7178 AAGGAGTGTGACCCCAATGAGGAGCAGCAGCAGCAACTTACGCAAAACAGCAAAAG 7237
DB 16414 AAGGAGTGTGACCCCAATGAGGAGCAGCAGCAGCAACTTACGCAAAACAGCAAAAG 16473
QY 7238 AATGATGGCAACAGAGTTCCTGCTACTTTCAGGAGCCTGCTTTTATCTTCTTATCTG 7297
DB 16474 AATGATGGCAACAGAGTTCCTGCTACTTTCAGGAGCCTGCTTTTATCTTATCTTATCTG 16533
QY 7298 ACCCCACCCACATCCTGCTGATTTGGCCCATTTTACAGACAGTGGATTGGTCCACTTACA 7357
```

```
DB 16534 ACCCCACCCACATCCTGCTGATTTGGCCCATTTTACAGACAGTGGATTGGTCCACTTACA 16593
QY 7358 GAGAGCTGATTTGGTGATTTTCAATCCCTGAGCTACACACAGAGTACTGATTTGGTATATT 7417
DB 16594 GAGAGCTGATTTGGTGATTTTCAATCCCTGAGCTACACACAGAGTACTGATTTGGTATATT 16653
QY 7418 TACAAACCTTGGAGCTAGACACAGAGTGTGAATGGTGTATTTTACAATCCCTTAGCTAGAC 7477
DB 16654 TACAAACCTTGGAGCTAGACACAGAGTGTGAATGGTGTATTTTACAATCCCTTAGCTAGAC 16713
QY 7478 ATAAAGTGTTCAGAGTCCCACTAGATTAGCTAGATAGATAGACAGAGACACTGATTT 7537
DB 16714 ATAAAGTGTTCAGAGTCCCACTAGATTAGCTAGATAGATAGACAGAGACACTGATTT 16773
QY 7538 GGTGGTGTTTACAAACCTTGGATTAGACACAGAGTGTGACTGTGTGTTTACAAACCTTGG 7597
DB 16774 GGTGGTGTTTACAAACCTTGGATTAGACACAGAGTGTGACTGTGTGTTTACAAACCTTGG 16833
QY 7598 AGCTAGACACAGAGTGTGATTTGGTGTATTTTACAATCTTTTAGCTAGAAATAAAGTTCC 7657
DB 16834 AGCTAGACACAGAGTGTGATTTGGTGTATTTTACAATCTTTTAGCTAGAAATAAAGTTCC 16893
QY 7658 CCAAGTCCCAACAGATTAGCTAGAT--AGAGTGTCTAATTTGGTGCATGCACGCCGA 7715
DB 16894 CCAAGTCCCAACAGATTAGCTAGATACAGAGTGTCTAATTTGGTGCATGCAC--AACCCGA 16952
QY 7716 GCTAGACACAGAGTGTGATTTGGTGCATATACAATCTCTGCTAGACATATAAAGTTCTC 7775
DB 16953 GCTAGACACAGAGTGTGATTTGGTGCATATACAATCTCTGCTAGACATATAAAGTTCTC 17012
QY 7776 CAAAGTCCCAACCTGACTCAGGAGGCCAGCAGCTTCGCCCTAGTGGATCC 7824
DB 17013 CAAAGTCCCAACCTGACTCAGGAGGCCAGCAGCTTCGCCCTAGTGGATCC 17060

RESULT 6
LOCUS AX394456 9432 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 1 from Patent WO0218638.
ACCESSION AX394456
VERSION AX394456.1 GI:21065594
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Risinger, C., Andersson, M.K., Lewander, T. and Olliasson, E.
TITLE Detection of cyp2d6 polymorphisms
JOURNAL Patent: WO 0218638-A 1 07-MAR-2002;
GEMINI Genomics PLC (GB)
FEATURES
Location/Qualifiers
source 1..9432
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 51.8%; Score 5325.6; DB 6; Length 9432;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 5895; Conservative 0; Mismatches 269; Indels 148; Gaps 26;

QY 637 TGTACAAAATACAGCCAGGCTAGTGTGGGCACTGTGTAATCCAGCTACTTGGAGGC 696
DB 186 TGAATAACAAAAGCTAGACGCTGTGGCAGACACCTGTGTAATCCAGCTACTTGGAGGC 245
QY 697 TGAGCAGAGGAATCGCTTGAACCTGGAGGTGGAGTTCAGTGCAGCGCAGATCGCATC 756
DB 246 TGAGCAGAGGAATTCCTTTGAAGCCTAGAGGTGAAGGTGTAGTGCAGCGCAGATTCATC 305
QY 757 ATTGCA-----CTCCACCTGGCGGAC-AGAGCAAACTCCGCTC----- 796
```


306	ATTGCCAATGAGGGAGCCACGCGCTGGGCNAAGAGGAAATCTCGTCTCCAAA	365
797	-----AAAAAATAATTTGGGCTCGGTGCTGCTAGTCCAGCTACTTGGGAG	849
366	AAAAAAAAAAAAAAAAAAGATTAGGCTGGTGGTGGCTAGTCCAGCTACTTGGGAG	425
850	GCAGGAGTCCACTTGATGTTGAGATTGCAGTGAAGCCATGATCCTGCACTGCACTCCGG	909
426	GCAGGGGTCCACTTGATGTCGAGACTGCACTGAGTGAAGCCATGCTGCACTGCACTCCGG	485
910	CCGGGCCAACAGAGTGAAGCCCTGCTAAAGAAAAATAAATAAAGCAACATATC	969
486	CCTGGGCACAGAGTGAAGCCCTGCTAAAGAA-----AAAAAATAAAGCAACATATC	540
970	CTAATAAAGGATCCTCCATAATGTTTCCACCAAGATTTCTAATCAGAAAATCGGAGGCCA	1029
541	CTGAACAAAGATCCTCCATAACGTTCCACCAAGATTTCTAATCAGAAAATCGGAGGCCA	600
1030	GGAAGCAGTGGAGAAATCAGACCCCTCAGGCAGCCCTGGAGGATGCTGTCAAGGCTGGGG	1089
601	GAAGCAGTGGAGAGGACGACCCCTCAGGCAGCCCGGGAGGATGTTGTCAAGGCTGGGG	660
1090	CAAGGCGCTTCAGGCTACCACTGGGAGCTCTGGGAACAGCCCTGTTGCAACAGAGT	1149
661	CAAGGCGCTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTTGCAACAGAGC	720
1150	CATGGCCGGCAGAGGCCA-GAATGTGGCTGAGCTGGGATCCATGTGACAGCTTTGAG	1208
721	CATAGCCCGGCAGAGCCAGGAATGTGGCTGGGCT-----757	
1209	GCTCACCGGAGCAGCCTCTTGGACAGGAGGTCCTATCCAGGAAACTCGGSCATGGCT	1268
758	-----GSGAGCAGCCTCTGGACAGAGTGCTCCATCCAGGAACTCCGSCATGGCT	810
1269	GGGAAGTGGGGTACTTGGTCGGGCTGTATGTGTGTGTGACTGTGTGTGTGAGAGAG	1328
811	GGGAAGTGGGGTACTTGGTCGGGCTGTATGTGTGTGTGACTGTGTGTGTGAGAGAG	870
1329	AATGTGTGCCCTGAGTGTGAGTGTGAGTGTGATGTGTGATATTGCTTTGTGTGGG	1388
871	AATGTGTGCCCTAAGTGTGAGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGGG	930
1389	TGATTTTCTGCATGTGAATCGTGTCCCTGCAAGTGTGAACAAGTGGACAAGTGTCTGGG	1448
931	TGATTTTCTGCGTGTGAATCGTGTCCCTGCAAGTGTGAACAAGTGGACAAGTGTCTGGG	990
1449	AGTGGACAAGAGATCTGTGCACCATCAGGTGTGTGTGATAGGCTGTGTGCATGTCAAGAT	1508
991	AGTGGACAAGAGATCTGTGCACCATCAGGTGTGTGTGATAGGCTGTGTGCATGTCAAGAT	1050
1509	GCAAGGTGAAGTGAAGGGACAGAGCCCATGATGCCACTCATCATCAGGAGCTCTAAGGCC	1568
1051	GCAAGGTGAAGTGAAGGGACAGAGCCCATGATGCCACTCATCATCAGGAGCTCTAAGGCC	1110
1569	CCAGGTAAGTGCCAGTGACATGAAGGCTGTGAAAGTGCATCTTGGAGTGGGAGGTTGGG	1628
1111	CCAGGTAAGTGCCAGTGACATGAAGGCTGTGAAAGTGCATCTTGGAGTGGGAGGTTGGG	1170
1629	GGTAGGAAAGGGCAAGGTCTATGTTCTGGAGGAGGGGTTGTGACTACATAGGTTGTATG	1688
1171	GGTAGGAAAGGGCAAGGCCATGTTCTGGAGGAGGGGTTGTGACTACATAGGTTGTATG	1230
1689	AGCCTAGCTGGGAGTGGATGGCCGGGCTCAGCTGAGACCTTGTTATCCAGAGCCTGT	1748
1231	AGCCTAGCTGGGAGTGGATGGCCGGGCTCAGCTGAAACCCCTGGTTATCCAGAGGCTTT	1290
1749	GTGGGCTTGGGGAGCTTGGAGTGGGAGAGGGGGTGACTTCTCCAGCAGGCTTTCTAC	1808
1291	GCAAGCTTCAGAGCTTGGAGTGGGAGAGGGGGTGACTTCTCCAGCAGGCTTTCTAC	1350
1809	CACCTTACCTTGGTTAAGGGCTTGGACAGGAGCAGCGCAAGGACCTCTGGAGCAGCC	1868
1351	CGGCTTACCTTGGTTAAGGGCTTGGACAGGAGCAGCGCAAGGACCTCTGGAGCAGCC	1410

QY	1869	CATACCTGCCCTGGCTGACTCTGCGCACTGGCGAGCAGTCAGTCAACACAGCAGGTTCACTCA	1928
DB	1411	CATACCCGCCCTGGCCTGACTCTGCCACTGGCAGCAGCAGTCAACACAGCAGGTTCACTCA	1470
QY	1929	CAGCAGAGGCGAAGGCCATCATCAGCTCCCTTTATAAGGGAAGGCTCACGCGCTCGGTG	1988
DB	1471	CAGCAGAGGCGAAGGCCCATCATCAGCTCCCTTTATAAGGGAAGGTCACGCGCTCGGTG	1530
QY	1989	TGCCGACAGTGTCTCTGCTCGTCTCTGTGCTGCTGGTGGGGTGGGGTGGCCAGTGTGTCC	2048
DB	1531	TGCTGAGTGTCTCTGCTGCTCTCTGTGCTGCTGGTGGGGTGGGGTGGCCAGTGTGTCC	1590
QY	2049	AGAGGACCCAGTTGTGTGAGGAGCCCATGAGGGCTAGAAAGCACTGTGTCCCTCGGCCA	2108
DB	1591	AGAGGACCCATTGTGTGTGAGGAGCCAGTATGGGGCTAGAAAGCACTGTGTCCCTCGGCCG	1650
QY	2109	TGATAGTGGGCATCTTCTGCTCTCTGGTGGACCTGATGCACCGGCAACCAACGCTGGGCTG	2168
DB	1651	TGATAGTGGGCATCTTCTGCTCTCTGGTGGACCTGATGCACCGGCGCAACGCTGGGCTG	1710
QY	2169	CACGCTACCCGCGCAGGTCCTCTGCCACTGCCCGGGCTGGCAACCTTGTGCTATGTGGAC	2228
DB	1711	CACGCTACCCACAGAGGCCCTCTGCCACTGCCCGGGCTGGCAACC-TGCTGCATGTGGAC	1769
QY	2229	TTCAGAAACACACCATACTCTCTTCGACCAAGGTGAGGAGGAGTCTCTGAGAGGCGGACAG	2288
DB	1770	TTCAGAAACACACCATACTCTCTTCGACCAAGGTGAGGAGGAGTCTCTGAGAGGCGGACAG	1829
QY	2289	GGTCTGAGGATGCCCAACACAGCAACCATGGTGTGGTGGTTAAACCAACAGGCTGGAT	2348
DB	1830	GGTCTGAGGCTCCCTTCACAGAAAGCAACCATGGATGGTGGTGGTAAACCAACAGGCTGGAC	1889
QY	2349	CAGAAGCCAGGCTGAGAAGGGAGCAGGTTTCGGGGACGTTCTCTGGGAAGGACATTTA	2408
DB	1890	CAGAAGCCAGGCTGAGAAGGGAGCAGGTTTCGGGGACG-TCTCGAAGAGGCAATTTA	1948
QY	2409	TACATGGCATGAAGGACTGGATTTTCCAAAGGCCAAGAGAGTAGGGCAAGGGCCTGGA	2468
DB	1949	TACATGGCATGAAGGACTGGATTTTCCAAAGGCCAAGAGAGTAGGGCAAGGGCCTGGA	2008
QY	2469	GGTGGAGCTGGACTTGTGGCAGTGGGCATGCAAGGCCATTGGGCAACATATGTTATGAGTA	2528
DB	2009	GGTGGAGCTGGACTTGGCAGTGGGCATGCAAGGCCATTGGGCAACATATGTTATGAGTA	2068
QY	2529	CAAAAGTCCCTTCTGCTGACACAGAGGAAGGCTTTGGGAATGGAAGATGAGTTAGTCC	2588
DB	2069	CAAAAGTCCCTTCTGCTGACACAGAGGAAGGCTTTGGGAATGGAAGATGAGTTAGTCC	2128
QY	2589	TGAGTGGCGTTTAAATCAGAAATCGAGGATGAAGGGGGTGCAGTGA CCGGTTCAAACC	2648
DB	2129	TGAGTGGCGTTTAAATCAGAAATCGAGGATGAAGGGGGTGCAGTGACCCGGTTCAAACC	2188
QY	2649	TTTTGCACTGTGGGTCTCTGGGCTCACTG-CTCAGCCGGATGGA CCAATCATCTGGGAAT	2707
DB	2189	TTTTGCACTGTGGGTCTCTGGGCTCACTGCTCCTCACCGCATGGA CCAATCATCTGGGAAT	2248
QY	2708	GGGATGCTAACTGGGGCTCTCGGCAATTTTGGTGA CTTTGCAGAGTCAATCCTGGGTG	2767
DB	2249	GGGATGCTAACTGGGGCTCTCGGCAATTTTGGTGA CTTTGCAGAGTCAATCCTGGGTG	2308
QY	2768	ACGCATCCAAACTGAGTTCTCCATCACAGAAAGTGTGACCCCAACCCCTGCCCCCAAGT	2827
DB	2309	ACGCATCCAAACTGAGTTCTCCATCACAGAAAGTGTGACCCCAACCCCGCCCAAGT	2368
QY	2828	CAGGAGCTGGGTCTCTCTTCCACTGCTCACTCTGTGTAGCCCGGGGGTGTGCTCAA	2887
DB	2369	CAGGAGCTGGGTCTCTCTTCCACTGCTCACTCTGTGTAGCCCGGGGGTGTGCTCAA	2428
QY	2888	GGTTCAATAGGACTAGGACTGTAGTCTGGGGTGATCCTGGCTTGACAGAGGCGCTGA	2947
DB	2429	GGTTCAATAGGACTAGGACTGTAGTCTGGGGTGATCCTGGCTTGACAGAGGCGCTGA	2488

4629 TGAACCTTTGCTGGGACACCCGGGGCTTCAAGCACAGGCTTGACAGGATCCTGTAAAGCCT 4688
QY
5128 AACTTCTCTCAACACACAGGAGGAGAGAGTGT-----CCCTGGGTGCTGAC 5174
Db
4689 GACCTCTCTCAACATAGGAGGCAAGAGAGTGTGAGGGCCGGGACCCCTGGGTGCTGAC 4748
QY
5175 CCATTGTGGGGACGCGATGCTGTCCAGTCCGCTGTCCAAACAGAGAGATCGACGAGTGTATAG 5234
Db
4749 CCATTGTGGGACGCGATGCTGTCCAGGCGGTGTCCNACAGAGATCGACGAGTGTATAG 4808
QY
5235 GCGAGGTGGGGACGCGATGCTGTGGGTGACAGGCTGTACATGCCCTTACACCACTGCGGTGA 5294
Db
4809 GCGAGGTGGGGACGCGATGCTGTGGGTGACAGGCTGTACATGCCCTTACACCACTGCGGTGA 4868
QY
5295 TTCAAGGTGGGACGCGCTTTGGGGACATCATCCCTGTGAGTGTGACCCATATGACATCCC 5354
Db
4869 TTCAAGGTGGGACGCGCTTTGGGGACATCATCCCTGTGAGTGTGACCCATATGACATCCC 4928
QY
5355 GTGACATCGAAGTACAGGCTTTCCGCATCCCTTAAGTAGGCTGTGGCCCTCTCACCCC 5414
Db
4929 GTGACATCGAAGTACAGGCTTTCCGCATCCCTTAAGTAGGCTGTGGCCCTCTCACCCC 4988
QY
5415 AGCTCAGCACGACGCTGTGATAGCCAGCATGGCTTACTGCGAGGTGGGCCCACTCT 5474
Db
4989 AGCTCAGCACGACGCTGTGATAGCCAGCATGGCTTACTGCGAGGTGGGCCCACTCT 5048
QY
5475 AGGAACCTGGGACGCTGTGATAGCCAGCATGGCTTACTGCGAGGTGGGCCCACTCT 5534
Db
5049 AGGAACCTGGGACGCTGTGATAGCCAGCATGGCTTACTGCGAGGTGGGCCCACTCT 5108
QY
5535 GGTCCAGAGTATAGGACGCTGTGATAGCCAGCATGGCTTACTGCGAGGTGGGCCCACTCT 5594
Db
5109 GGTCCAGAGTATAGGACGCTGTGATAGCCAGCATGGCTTACTGCGAGGTGGGCCCACTCT 5168
QY
5595 AACACGACCTCCGAGATGTTGGAGGACCCAGCCCTGACAGGAGAGGGGACGCTGTGG 5654
Db
5169 AACACGACCTCCGAGATGTTGGAGGACCCAGCCCTGACAGGAGAGGGGACGCTGTGG 5228
QY
5655 GTGCTCTGAGAGGTGTGATGCGCCCTGCTGTGGGTGCGAGAGGTGACTGTGAGCTTT 5714
Db
5229 GTGCTCTGAGAGGTGTGATGCGCCCTGCTGTGGGTGCGAGAGGTGACTGTGAGCTTT 5288
QY
5715 CTCGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCGAGGAGTGTGTCCCGCT 5774
Db
5289 CTCGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCGAGGAGTGTGTCCCGCT 5348
QY
5775 GTGTTTGTGGCAGGGGTCCTGAGCATCTCTAGAGTCCAGTCCCACTCTCACTCCATCT 5834
Db
5349 GTGTTTGTGGCAGGGGTCCTGAGCATCTCTAGAGTCCAGTCCCACTCTCACTCCATCT 5408
QY
5835 CCTGCCAGGGAACGACACTGATCAACCACTGTCTATCGGTCTGAAGATGAGCGGCTC 5894
Db
5409 CTGCGCCAGGGAACGACACTGATCAACCACTGTCTATCGGTCTGAAGATGAGCGGCTC 5468
QY
5895 TCGGAGAACCTTCCGCTTCCACCCCGAACAATCTCTGATGTCGAGGAGGAGTGTGTG 5954
Db
5469 TCGGAGAACCTTCCGCTTCCACCCCGAACAATCTCTGATGTCGAGGAGGAGTGTGTG 5528
QY
5955 AAGCCGAGGCTTCTGCTGCTTCTCAGCAGGTGCTGTGGGAGCCCGGCTCCCTGTCC 6014
Db
5529 AAGCCGAGGCTTCTGCTGCTTCTCAGCAGGTGCTGTGGGAGCCCGGCTCCCTGTCC 5588
QY
6015 CTTTCCGTGGAGTCTTGAGGGGTATACCCAGGAGCCAGGCTCTAGCGCCCTCCGC 6074
Db
5589 CTTTCCGTGGAGTCTTGAGGGGTATACCCAGGAGCCAGGCTCTAGCGCCCTCCGC 5648
QY
6075 TCCCAACAGGCGCGTGTGATGCTTCCGAGGAGCCCTGCGGAGTGTGAGCTTCTCTCT 6134
Db
5649 TCCCAACAGGCGCGTGTGATGCTTCCGAGGAGCCCTGCGGAGTGTGAGCTTCTCTCT 5708
QY
6135 TCTTCACTCTCTGTGAGCACTTTCAGCTTCTCGTGGCGCGGAGCAGCGCCGCGCCA 6194
Db
5709 TCTTCACTCTCTGTGAGCACTTTCAGCTTCTCGTGGCGCGGAGCAGCGCCGCGCCA 5768

QY 6195 GCCACTCTGCTGCTCAGCTTTCTGGTGAGCCCATCCCTACGAGCTTTTGTGCTGTC 6254
Db
5769 GCCACCATGGTGTCTTCTGTTCTGGTGAGCCATCCCTATGAGCTTTGTGCTGTC 5828
QY 6255 CCCTAGATAAGGCTACCTAGTCCCAAGCTGTCTCTAGCCAGAGCTCTTAATGTACA 6314
Db
5829 CCCTAGATAAGGCTACCTAGTCCCAAGCTGTCTCTAGCCAGAGCTCTTAATGTACA 5888
QY 6315 ATAAAGCAATGTGGTAGTTCCTAGTCCCAAGCTGTCTCTAGCCATCTCTGGGATCATCC 6374
Db
5889 ATAAAGCAATGTGGTAGTTCCTAGTCCCAAGCTGTCTCTAGCCATCTCTGGGATCATCC 5948
QY 6375 TCCTCAGGGAACCCACCTGCTCTCTCTGCTTACCCACCCGCTGCGCGCATTTG 6434
Db
5949 TCCTCAGGGAACCCACCTGCTCTCTCTGCTTACCCACCCGCTGCGCGCATTTG 6008
QY 6435 AGAC--GGGTACGTTGAGGCTGAGCAGATGTCTGCTTACCTTGGCCATAATCCCATGTCC 6492
Db
6009 AGACAGGAGTACGTTGAGGCTGAGCAGATGTCTGCTTACCTTGGCCATAATCCCATGTCC 6068
QY 6493 CCCACTGACCCAACTCTGATGCTCCAGATTGGTGACAAGGACTACATTGTCTTGGCATGT 6552
Db
6069 CCCACTGACCCAACTCTGATGCTCCAGATTGGTGACAAGGACTACATTGTCTTGGCATGT 6128
QY 6553 GGGGAAGGGCCAGAAATGGGCTGACTAGAGGTGTGAGTGTGAGTGTGGTGGAGAG 6612
Db
6129 GGGGAAGGGCCAGAAATGGGCTGACTAGAGGTGTGAGTGTGAGTGTGGTGGAGAG 6188
QY 6613 GCGAGGACTCAGCCTGGAGGCGCATATTTCAAGGCTTAATCTCAGCCCACTCCCATCAGGG 6672
Db
6189 GCGAGGACTCAGCCTGGAGGCGCATATTTCAAGGCTTAATCTCAGCCCACTCCCATCAGGG 6248
QY 6673 ACAGAGTCTTCCAGCAACCATCAACAGTCACTCTCTCATATATGACACCCCAAAA 6732
Db
6249 ACAGAGTCTTCCAGCAACCATCAACAGTCACTCTCTCATATATGACACCCCAAAA 6308
QY 6733 TGGGAGACAAATCAT---GTCAGGAGCTATATGCGAGG---CTACTCCAGGGCTC 6785
Db
6309 CGGAAGACAAATCATGCGGTGAGGAGCTATATGCGAGGCTACTCTACTCCAGGGCTC 6368
QY 6786 AGTGGCAGGTGCGCAGAACATTCTCCGTGGAGGCGCCAGGAAACCCAGGACCGAGCCAC 6845
Db
6369 AGTGGCAGGTGCGCAGAACATTCTCCGTGGAGGCGCCAGGAGCCAGGACTGAGCCAC 6428
QY 6846 GCGCTCAGCCT 6857
Db 6429 CACCTCAGCCT 6440

RESULT 7

AX687027
LOCUS
DEFINITION
Sequence 1 from Patent EP1281755.
ACCESSION
AX687027
VERSION
AX687027.1 GI:29409531
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Milos P.M. and Webb, S.M.
TITLE
Variants of the human cyp2d6 gene
JOURNAL
Patent: EP 1281755-A 1 05-FEB-2003;
Pfizer Products Inc. (US)
FEATURES
Location/Qualifiers
1. 9432
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

linear

PAT 31-MAR-2003

Query Match		51.8%;	Score 5325.6;	DB 6;	Length 9432;
Best Local Similarity		93.4%;	Pred. No. 0;		
Matches 5895;		Conservative 0;	Mismatches 269;	Indels 148;	Gaps 26;
QY	637	TGTACAAAATCAGCCAGGATGTTGGTGGGACCTGTAAATCCACAGCTACTTGGGAGGC	696		
DB	186	TGAATAACAAAAGCTAGACGTGGTGGACACACCTGTAAATCCACAGCTACTTGGAGGC	245		
QY	697	TGAGGAGGAGAAATCGCTTGAACCTCTGGAGGTGGAGTTGCAGTGAGCCGAGATCGCATC	756		
DB	246	TGAGGAGGAGAAATTCCTGAAGCCTAGAGTTGAAGTGTAGTGAGCCGAGATTGCATC	305		
QY	757	ATTGCA-----CTCCACCTGGGGAC-AGAGCAAAATCTCGTCTC-----	796		
DB	306	ATTGCAATGGAGGGAGCCACAGCCTGGGCAACAAGAGGAAATCTCCGTCTCAAAA	365		
QY	797	-----AAAAAAAATATTGGGCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGGAG	849		
DB	366	AAAAAAAATTAATTTGGGCTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGGAG	425		
QY	850	GCAGGAGGTCCTTGTATGTTGAGATTGCAAGTGAGCCATGATCCTGCCACTGCACCTCCGG	909		
DB	426	GCAGGGGTCCTTGTATGTCGAGACTGCAGTGAGCCATGATCCTGCCACTGCACCTCCGG	485		
QY	910	CCCGGCAACAGAGTGAGACCTGTCTAAAGAAAAAATAAATAAATAAAGCAATATC	969		
DB	486	CCTGGGCAACAGAGTGAGACCTGTCTAAAGAA-----AAAAAAAATAAAGCAATATC	540		
QY	970	CTAAATAAAGGATCTCCATAATGTTTCCACAGATTTCTAATCAGAAACATGGAGCCA	1029		
DB	541	CTGAACAAGGATCTCCATAAGCTTCCACACAGATTTCTAATCAGAAACATGGAGCCA	600		
QY	1030	GGAGCAGTGGGAATGACGACCTCAGGCAGCCCTGGAGGATGCTGTACAGGCTGGGG	1089		
DB	601	GAAAGCAGTGGAGGAGGACGACCTCAGGCAGCCCGGGAGGATGTTGTACAGGCTGGGG	660		
QY	1090	CAAGGCGCTTCAGGCTACCAACTGGGAGCTCTGGGAAAGCCCTGTTTGAACAGGAATG	1149		
DB	661	CAAGGCGCTTCCGGCTACCAACTGGGAGCTCTGGGAAAGCCCTGTTTGAACAGGAATG	720		
QY	1150	CATGGCCCGCCAGAGCCCA-GAATGTGGGCTGAGCTGGGATCCATGTACACGCTTTGAG	1208		
DB	721	CATAGCCCGGCCAGAGCCAGGAATGTGGGCTGGGCT-----	757		
QY	1209	GCTCACCGGAGCAGGCTCTGACAGGAGAGGTCCCATCCAGGAAACCTCGGCGATGGCT	1268		
DB	758	-----GGGAGCAGCCTCTGACAGAGAGTGTGTCCTATCCAGGAAACCTCCGCGATGGCT	810		
QY	1269	GGAAATGGGGTACTTGGTCCGGGTCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1328		
DB	811	GGAAATGGGGTACTTGGTCCGGGTCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	870		
QY	1329	AATGTGTCCCTGAGTGTGAGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1388		
DB	871	AATGTGTCCCTGAGTGTGAGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	930		
QY	1389	TGATTTTCTGCATGTGTAAATCGTGTCCCTGCAAGTGTGAACAGTGCACAAGTGTCTGGG	1448		
DB	931	TGATTTTCTGCATGTGTAAATCGTGTCCCTGCAAGTGTGAACAGTGCACAAGTGTCTGGG	990		
QY	1449	AGTGGACAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTGTGTGTGTGTGTGTGTGT	1508		
DB	991	AGTGGACAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTGTGTGTGTGTGTGTGTGT	1050		
QY	1509	GCAGGTGAGTGAAGGAGCCAGGCCCATGATGCCACTCATCATCAGGAGCTCTAAGGCC	1568		
DB	1051	GCAAGGTGAAGTGAAGGAGCCAGGCCCATGATGCCACTCATCATCAGGAGCTCTAAGGCC	1110		
QY	1569	CCAGGTGAAGTGCAGGTGACAGATAAGGGTGTCAAGGTCACTCTGGAAGTGGGAGGTGGG	1628		
DB	1111	CCAGGTGAAGTGCAGGTGACAGATAAGGGTGTCAAGGTCACTCTGGAAGTGGGAGGTGGG	1170		
QY	1629	GGTAGGGAAGGCAAGGTCATGTTCTGGAGGAGGGGTTGTGACTATACATTAGGGTGTATG	1688		

DB	1171	CGTAGGAAAGGCGAAGGCCATGTTCTGGAGGAGGGTTGTGACTACATTAGGGTGTATG	1230		
QY	1689	AGCCTAGCTGGGAGGTGGATGCCCGGTCCTACTGAGACCTCGTTATCCAGAGGCTGT	1748		
DB	1231	AGCCTAGCTGGGAGGTGGATGCCCGGTCCTACTGAAACCTCGTTATCCAGAGGCTTT	1290		
QY	1749	GTGGGCTGGGGAGCTTGGAGTGGGAGAGGGGGTGACTTCTCCGACAGGCTTTCTAC	1808		
DB	1291	GCAGGCTTCAGGAGCTTGGAGTGGGAGAGGGGGTGACTTCTCCGACAGGCCCCCACC	1350		
QY	1809	CACCTTACCTCGGTAAAGGCTTGGAGCAGGAAGCGGCAAGGACCTCTGGAGCAGCC	1868		
DB	1351	CGGCCCTTACCTGGTAAAGGCTTGGAGCAGGAAGCGGGCAAGAACCTCTGGAGCAGCC	1410		
QY	1869	CATACCTGCCCTGGCTGACTCTGCCACTGGCAGCAGCTCAACACAGCAGGTTCACTCA	1928		
DB	1411	CATACCGGCTTGGCTGACTCTGCCACTGGCAGCAGCTCAACACAGCAGGTTCACTCA	1470		
QY	1929	CAGCAGAGGGCGAAGGCCATCATAGCTCCCTTTTAAAGGGAAGGCTCAACGCTCGGTG	1988		
DB	1471	CAGCAGAGGGCAAGGCCATCATAGCTCCCTTTTAAAGGGAAGGCTCAACGCTCGGTG	1530		
QY	1989	TGCCAGAGGTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2048		
DB	1531	TGCTGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1590		
QY	2049	AGAGAGCCAGTGTGTAGTGAGGAGCCATGCGGGCTAGAAGCAGCTGTGCTGCTGCTG	2108		
DB	1591	AGAGAGCCCATTTGTGTAGTGAGGAGGTATGGGGCTAGAAGCAGCTGTGCTGCTGCTG	1650		
QY	2109	TGATAGTGGCCATCTTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2168		
DB	1651	TGATAGTGGCCATCTTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1710		
QY	2169	CACGCTACCCGCGCAGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2228		
DB	1711	CACGCTACCCGCGCAGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1769		
QY	2229	TTCCAGAACACACACTACTGCTTCCAGCCAGGTTGAGGAGGAGGCTCTGAGGCGGCG	2288		
DB	1770	TTCCAGAACACACACTACTGCTTCCAGCCAGGTTGAGGAGGAGGCTCTGAGGCGGCG	1829		
QY	2289	GGTCTCTGAGGATGCCCGACACACAGCAAAATGCTGGTGGTGGTGGTGGTGGTGGTGG	2348		
DB	1830	GGTCTCTGAGGCTCCCTTACCAGAACAAATGATGGTGGTGGTGGTGGTGGTGGTGG	1889		
QY	2349	CAGAAAGCAGGCTGAGAGAGGGGAAAGAGTTTGGGGGAGCTTCTGGGGGAGGACATTTA	2408		
DB	1890	CAGAAAGCAGGCTGAGAGAGGGGAAAGAGTTTGGGGGAGGAGGCTTCTGGGGGAGG	1948		
QY	2409	TACATGGCATGAAGGATGGATTTTCCAAAGCCCAAGGAGAGTGGGCAAGGCTGGA	2468		
DB	1949	TACATGGCATGAAGGATGGATTTTCCAAAGCCCAAGGAGAGTGGGCAAGGCTGGA	2008		
QY	2469	GGTGGAGCTGGAGTGGGAGTGGGATGCAAGCCCATTTGGGCAACATATGTTATGAGTA	2528		
DB	2009	GGTGGAGCTGGAGTGGGAGTGGGATGCAAGCCCATTTGGGCAACATATGTTATGAGTA	2068		
QY	2529	CAAGTCTCTTCTGCTGACACAGAAAGGAAAGGCTTGGGAATGGAAGATGATTTAGTTC	2588		
DB	2069	CAAGTCTCTTCTGCTGACACAGAAAGGAAAGGCTTGGGAATGGAAGATGATTTAGTTC	2128		
QY	2589	TGAGTGGCTTTAAATCAGGAATCAGGAATGAAGGGGTGAGTGGGCTTCAAGCC	2648		
DB	2129	TGAGTGGCTTTAAATCAGGAATCAGGAATGAAGGGGTGAGTGGGCTTCAAGCC	2188		
QY	2649	TTTTGCACTGTGGGCTCTCGGGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2707		
DB	2189	TTTTGCACTGTGGGCTCTCGGGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2248		
QY	2708	GGGATGCTAACTGGGCTCTCGGCAATTTTGGTGAATCTTTCAGAGGTCAATCTGGGTG	2767		

4888 GCTTCAATGATGAGAACCTGCGCATAGTGGTGGGTAACTGTTCTTCTGCGGATGGTGA 4947
|||||
4449 GCTTCAATGATGAGAACCTGCGCATAGTGGTGGTGAACCTGTTCTTCTGCGGATGGTGA 4508
|||||
4948 CCACTCGACACGCTGGCTGGGGCTCTCTGCTCATGATCCTACACCTGGATGTGCAGC 5007
|||||
4509 CCACTCGACACGCTGGCTGGGGCTCTCTGCTCATGATCCTACATCCGGATGTGCAGC 4568
|||||
5008 GTGAGCCAGCTGGGGCCCAAGCGAGGGAATGAGGGAAGGATACAGTGGGGGCCCC 5067
|||||
4569 GTGAGCCCACTCTGGGAAAAGTGCAGGGGCGGAGGAGGAAGGATACAGCGGGGCCCA 4628
|||||
5068 TGGGCTTAGCTGGGACACCGGGGCTTCCAGCACAGCGTGGCGAGGCTCTGTAAAGCTT 5127
|||||
4629 TGAACCTTGTGGGACACCGGGGCTTCAAGCACAGGCTTGACAGGATCTGTAAAGCTT 4688
|||||
5128 AACTTCTCTCAACACAGGAGGAAGGAGTGT-----CCCTGGGTGTGAC 5174
|||||
4689 GACCTCTTCAACATAGGAGGGAAGAGGATGTGAGGGCGGACCCCTGGGTGTGAC 4748
|||||
5175 CCATTGTGGGACGCACTGTCTGTCCAGTCCGTGTCCAAAGAGGATCGACGAGTGAATAG 5234
|||||
4749 CCATTGTGGGACGCACTGTCTGTCCAGGCGGTGTCCAAAGAGGATCGACGAGTGAATAG 4808
|||||
5235 GGCAGTGGCGGACACAGAGATGGGTGAACAGGCTCACATGCCCTTACACACTGCCGTGA 5294
|||||
4809 GGCAGTGGCGGACACAGAGATGGGTGAACAGGCTCACATGCCCTTACACACTGCCGTGA 4868
|||||
5295 TTCACAGGTGGACGCGCTTGGGACATCATCCCCCTGAGTGTGACCCATATGACATCCC 5354
|||||
4869 TTCAAGGTGGACGCGCTTGGGGAATCGTCCCCCTGGGTGTGACCCATATGACATCCC 4928
|||||
5355 GTGACATCGAAGTACAGGCTTCCGCACTCCCTAAGGTAGGCTTGGCGCCCTCTCACCCC 5414
|||||
4929 GTGACATCGAAGTACAGGCTTCCGCACTCCCTAAGGTAGGCTTGGCGCCCTCTCACCCC 4988
|||||
5415 AGCTCAGACACGACCTGGTGTATGATCCCGAGATGGCTACTGTCAGGTGGGCCCACTCT 5474
|||||
4989 AGCTCAGACACGACCTGGTGTATGATCCCGAGATGGCTACTGTCAGGTGGGCCCACTCT 5048
|||||
5475 AGGACCTTGGCCACCTAGTCTCAATGCAACCACTGACTGTCCCACTTGGGTGGGG 5534
|||||
5049 AGGAACTTGGCCACCTAGTCTCAATGCAACCACTGACTGTCCCACTTGGGTGGGG 5108
|||||
5535 GGTCCAGAGATATAGGACAGGCTGGCCCTGTCCATCCAGAGCCCGCTAGTGGGAGACA 5594
|||||
5109 GGTCCAGAGATATAGGACAGGCTGGCCCTGTCCATCCAGAGCCCGCTAGTGGGAGACA 5168
|||||
5595 AACCCAGGACCTGCAGAAATGTTGGAGACCCAGCGCTTGCAGGAGAGGGGCGAGTGG 5654
|||||
5169 AACCCAGGACCTGCAGAAATGTTGGAGGACCCAGCGCTTGCAGGAGAGGGGCGAGTGG 5228
|||||
5655 GTGCTCTGAGAGGTGACTGGCCCTGCTGTGGGGTGGAGAGGGTACTGTGAGCTTT 5714
|||||
5229 GTGCTCTGAGAGGTGACTGGCCCTGCTGTGGGGTGGAGAGGGTACTGTGAGCTTT 5288
|||||
5715 CTCGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAGGAGTGTGTCCCGCT 5774
|||||
5289 CTCGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAGGAGTGTGTGTCCCGCT 5348
|||||
5775 GTGTTTGGTGGCAGGGGTCCTCAGCATCTCAGTCCAGTCCCACTCTCACCTCGCATCT 5834
|||||
5349 GTGTTTGGTGGCAGGGGTCCTCAGCATCTCAGTCCAGTCCCACTCTCACCTCGCATCT 5408
|||||
5835 CTGTCACGAGGAACGACATCAATCAACCACTGTATCGGTGCTGAAGGATGAGCCGCTC 5894
|||||
5409 CTGTCACGAGGAACGACATCAATCAACCACTGTATCGGTGCTGAAGGATGAGCCGCTC 5468
|||||
5895 TGGGAGAACCTTCCGCTTCCACCCCGAACAATCTCTGGATGCCAGGCGCACTTTGTG 5954
|||||
5469 TGGGAGAACCTTCCGCTTCCACCCCGAACAATCTCTGGATGCCAGGCGCACTTTGTG 5528
|||||
5955 AAGCCGGAGGCTTCTCTGCTTTCTCAGCAGGTGCTGTGGGGAGCCCGGCTCCCTGTGCC 6014
|||||

5529 AAGCCGGAGGCTTCTCTGCTTTCTCAGCAGGTGCTGTGGGAGCCCGCTCCCTGTGCC 5588
|||||
6015 CTTTCGTGAGTCTTTGACAGGGGTATCACCCAGGAGCCAGGCTCACTGACGCCCTTCCCC 6074
|||||
5589 CTTTCGTGAGTCTTTGACAGGGGTATCACCCAGGAGCCAGGCTCACTGACGCCCTTCCCC 5648
|||||
6075 TCCCAACAGCCCGCGTGCATGCTCGGGAGCCCTGCGCCGCGCATGAGAGCTTCTCTCT 6134
|||||
5649 TCCCAACAGCCCGCGTGCATGCTCGGGAGCCCGCTGCGCCGCGCATGAGAGCTTCTCTCT 5708
|||||
6135 TCTTCACTCTCTGCTGACAGCTTCTGCTGAGCTTCTCGTGGCGCCGCGACGCCGCCCA 6194
|||||
5709 TCTTCACTCTCTGCTGACAGCTTCTGCTGAGCTTCTCGTGGCGCCGCGACGCCGCCCA 5768
|||||
6195 GGCACCTCTGCTGCTGCTGAGCTTCTGCTGAGCTTCTCGTGGCGCCGCGACGCCGCCCA 6254
|||||
5769 GGCACCATGCTGCTTCTGCTGAGCTTCTCGTGGCGCCGCGACGCCGCCCA 5828
|||||
6255 CCGCTAGAAATGGGTGACCTAGTCCCGAGCTGCTCCCTAGCCAGAGGCTCTAAATGACA 6314
|||||
5829 CCGCTAGAAATGGGTGACCTAGTCCCGAGCTGCTCCCTAGCCAGAGGCTCTAAATGACA 5888
|||||
6315 ATAAAGCAATGTGTAGTTCCTAACTTGGGTGCTGCTCAAGCCCTCTGTTGGGATCATCC 6374
|||||
5889 ATAAAGCAATGTGTAGTTCCTAACTTGGGTGCTGCTCAAGCCCTCTGTTGGGATCATCC 5948
|||||
6375 TCCTCAGGGCAACCCCACTGCTGCTCATCTGCTTACCCCAAGCCCTGCGCGCATTTG 6434
|||||
5949 TCCTCAGGGCAACCCCACTGCTGCTCATCTGCTTACCCCAAGCCCTGCGCGCATTTG 6008
|||||
6435 AGAC--GGGTAGCTTGGGCTGAGCAGATGTGAGTTCAGTTCACCTTTGCCATAATCCCATGTCC 6492
|||||
6009 AGCAGGGGTAGCTTGGGCTGAGCAGATGTGAGTTCAGTTCACCTTTGCCATAATCCCATGTCC 6068
|||||
6493 CCCACTGACCCCACTTGAATGCTGCCAGATTTGGTGAAGGACTACATTTCTCTGGCATGT 6552
|||||
6069 CCCACTGACCCCACTTGAATGCTGCCAGATTTGGTGAAGGACTACATTTCTCTGGCATGT 6128
|||||
6553 GGGGAAGGGCCAGAAATGGGCTGAGTGGGCTGAGTGGGCTGAGTGGGCTGAGTGGTGGAGAG 6612
|||||
6129 GGGGAAGGGCCAGAAATGGGCTGAGTGGGCTGAGTGGGCTGAGTGGGCTGAGTGGTGGAGAG 6188
|||||
6613 GGCAGGACTCAGCTGGAGGCCCATATTTTACGGCTTAACTCAGCCACCCCAATCATCAGG 6672
|||||
6189 GGCAGGACTCAGCTGGAGGCCCATATTTTACGGCTTAACTCAGCCACCCCAATCATCAGG 6248
|||||
6673 ACAGAGCTCTGCTCCAGCACCATCAACAGTCACTTCCCTTCCATATATGACACCCCAAAA 6732
|||||
6249 ACAGAGCTCTGCTCCAGCACCATCAACAGTCACTTCCCTTCCATATATGACACCCCAAAA 6308
|||||
6733 TGGGAAGACAATCAT---GTGAGGAGCTATATGCGAGG---CTACTCCAGGGCTC 6785
|||||
6309 CGGAAGACAAATCATTGGGCTGAGGAGCTATATGCGAGGCTTACCTTACCTCCAGGGCTC 6368
|||||
6786 AGTCGGCAGGTGCGCAAGACATTTCCCTGGGAGGCCCGCCAGGAAAACCCAGGACCGAGCCAC 6845
|||||
6369 AGTCGGCAGGTGCGCAAGACATTTCCCTGGGAGGCCCGCCAGGAAAACCCAGGACCGAGCCAC 6428
|||||
6846 GCGCCCTCAGCCTT 6857
|||||
6429 CACCCCTCAGCCT 6440
|||||

RESULT 8

HUMCYP2D6

LOCUS

Human cytochrome

P450 IID6 (CYP2D6)

DEFINITION

M33388

ACCESSION

M33388.1

VERSION

GI:181303

KEYWORDS

cytochrome P450; cytochrome P450 IID6.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

9432 bp DNA linear PRI 22-NOV-1994
complete cds.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 9432)
Kimura,S., Umeno,M., Skoda,R.C., Meyer,U.A. and Gonzalez,F.J.
The human debrisoquine 4-hydroxylase (CYP2D6) locus: sequence and
identification of the polymorphic CYP2D6 gene, a related gene, and
a pseudogene
Am. J. Hum. Genet. 45 (6), 889-904 (1989)
2574001
90072069
PUBMED
COMMENT
Original source text: Human DNA, clone lambda2D-18/2.
Draft entry and computer-readable sequence for [Am. J. Hum. Genet.
45, 889-904 (1989)] kindly submitted
by S.Kimura, 29-MAR-1990

FEATURES
source

Location/Qualifiers
1..9432
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="22q13.1"
Join(1532..1799,2503..2674,3225..3377,3466..3626,
4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)
/gene="CYP2D6"
Join(1532..1799,2503..2674,3225..3377,3466..3626,
4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)
/gene="CYP2D6"
/product="cytochrome P450 IID6"
/note="G00-132-127"
1532..1799
/gene="CYP2D6"
/note="cytochrome P450 IID6; G00-132-127"
/number=1
Join(1620..1799,2503..2674,3225..3377,3466..3626,
4060..4236,4427..4568,4776..4963,5418..5559,5658..5836)
/gene="CYP2D6"
/codon_start=1
/product="cytochrome P450 IID6"
/protein_id="AA53500.1"
/db_xref="GI:181304"
/db_xref="GDB:G00-132-127"
/translation="MGLEALVPLAVIAIFLLVLMHRRORWAARYPPGPIPLPLGLG
NLLHVDFTNYPCFQDLRRRGDFVSLQAWTPVVLNGLAAVREALVTHGEDTADRP
PVPVPIQLIGFGRSGVFLARYGPWRQRRESVSTENLGLGKSLQWTEEAACL
CAAFANSGRPRLNGLLDKAVNSVIASTLCGRFEYDDPRFLRLDLQAQGLKEEG
FLREVLNAPVLLHLPALGAKVLRQKAPLTLQDELLTEHRMTWDPAPQPRDLTEAFL
AMEKAGNPSSFNENLRIVYADLFSAGVVTSTTLAWGLLMLHPDVQRVQOE
IDDI7GVRRPMGQAHMPYTTAVIHEVQRFQDIVPLGVTHMTSRDIEVOGFRIPKG
TTLTNLSVLKDEAVWEKPRFHPHEHFLDAQGHFVKPEAFIPRSAGRRACLGLEPLAR
MELFLFTSLLOHFSFSVPTGQPRSHHGVAFLVSPSPYELCAVPR"
1800..2502
/gene="CYP2D6"
/note="G00-132-127"
/number=1
2503..2674
/gene="CYP2D6"
/note="G00-132-127"
/number=2
2675..3224
/gene="CYP2D6"
/note="G00-132-127; does not fit consensus"
/number=2
3225..3377
/gene="CYP2D6"
/note="G00-132-127"
/number=3
3378..3465
/gene="CYP2D6"
/note="G00-132-127"
/number=3
3466..3626
/gene="CYP2D6"
/note="G00-132-127"
/number=4

intron
exon
intron
exon
intron
exon
intron
exon
intron
exon
ORIGIN
Query Match 51.8%; Score 5325.6; DB 9; Length 9432;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 5895; Conservative 0; Mismatches 269; Indels 148; Gaps 26;
QY 637 TGTCAAAAAATCAGCCAGGCGATGGTGGTGGGCACCTGTAATCCAGCTACTTTGGAGGC 696
Db 186 TGAATAACAAAAGCTAGACGTGTGGCACACACCTGTAATCCAGCTACTTAGAGGC 245
QY 697 TGAGGAGGAGAAATCGCTTGAACCCCTGGAGGTGGAGTTGCACTGAGCCGAGATCGCATC 756
Db 246 TGAGGAGGAGAAATGCTTGAAGCCTAGAGGTGAAGTGTGTAGTGAGCCGAGATTGCATC 305
QY 757 ATTGCA-----CTCCCACTGGGGGAC-AGACAAACCTCCGCTC-----796
Db 306 ATTGCACAATGGAGGGGAGCCACCAGCCTGGGCAACAAGAGAAATCTCCGCTTCAAAA 365
QY 797 -----AAAAAAAATTTATTTGGGCTCGGTGGTGGCTGTAGTCCCACTACTTTGGGAG 849
Db 366 AAAAAAAAATAAAGAAATTTAGGCTGGTGGTGGCTGTAGTCCCACTACTTTGGGAG 425
QY 850 GCAGAGGTCCACTTGATTTGAGATTGCAGTGAGCCATGATCTTGCACCTGCATCCGG 909
Db 426 GCAGGGGTCCACTTGATGTGAGACTGCAGTGAGCCATGATCTTGCACCTGCATCCGG 485
QY 910 CCGGGCAACAGAGTGAGACCCCTGCTTAAAGAAAAATAAATAAAGAAACACATATC 969
Db 486 CTTGGCAACAGAGTGAGACCCCTGCTTAAAGAA-----AAAAAATAAAGCAACATATC 540
QY 970 CTAATAAGGATCTCCATATATTTTCCACAGATTTCTAATCAGAAACATGGAGGCCA 1029
Db 541 CTGAACAAGGATCTCTCATAAGCTTCCACCAAGATTTCTAATCAGAAACATGGAGGCCA 600

5355 GTGACATCGAAGTACAGGGCTTCGGCATCCCTAAGGTAGGCTTGGCGCCCTCCTCAACCC 5414
5429 GTGACATCGAAGTACAGGGCTTCGGCATCCCTAAGGTAGGCTTGGCGCCCTCCTCAACCC 4988
5415 AGCTCAGCACCAGCACTGGTGTATAGCCAGCATGGCTACTGCGAGGTGGGCCACTCT 5474
4989 AGCTCAGCACCAGCACTGGTGTATAGCCAGCATGGCTACTGCGAGGTGGGCCACTCT 5048
5475 AGGAACCTTGGCCACCTAGTCCCTCAATGCCACACACTGACTGTCCCCACTTGGGTGGG 5534
5049 AGGAACCTTGGCCACCTAGTCCCTCAATGCCACACACTGACTGTCCCCACTTGGGTGGG 5108
5535 GGTCCAGAGTATAGGAGGCTGGCTGTGTCATCCAGAGCCCCCGTCTAGTGGGAGACA 5594
5109 GGTCCAGAGTATAGGAGGCTGGCTGTGTCATCCAGAGCCCCCGTCTAGTGGGAGACA 5168
5595 AACCAGGACCTGCAGAAATGTTGGAGACCCAGCGCTGCGAGGAGAGGGGAGGTGG 5654
5169 AACCAGGACCTGCAGAAATGTTGGAGACCCAGCGCTGCGAGGAGAGGGGAGGTGG 5228
5655 GTGCTCTGAGAGGTGTGACTGGCCCTGCTGTGGGGTGGAGAGGCTACTGTGGAGCTT 5714
5229 GTGCTCTGAGAGGTGTGACTGGCCCTGCTGTGGGGTGGAGAGGCTACTGTGGAGCTT 5288
5715 CTGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAGGAGTGTGTGCCCGT 5774
5289 CTGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAGGAGTGTGTGCCCGT 5348
5775 GTGTTTGGTGGCAGGGGTCCAGCATCTTAGAGTCCAGTCCCGACTCTCACCTGATCT 5834
5349 GTGTTTGGTGGCAGGGGTCCAGCATCTTAGAGTCCAGTCCCGACTCTCACCTGATCT 5408
5835 CTGGCCAGGGAACGACACTCATCAACACTGTCTCATCGTGTGAAGATGAGGCCGTC 5894
5409 CTGGCCAGGGAACGACACTCATCAACACTGTCTCATCGTGTGAAGATGAGGCCGTC 5468
5895 TGGGAGAAGCCCTTCGCTTCCACCCCGAACACTTCTCGATGTCGCCAGGCGCACTTTGTG 5954
5469 TGGGAGAAGCCCTTCGCTTCCACCCCGAACACTTCTCGATGTCGCCAGGCGCACTTTGTG 5528
5955 AAGCCGAGGCTTCTGCTTCTCAGCAGTGTCTGCGGAGCCCGGCTCCCTGTCC 6014
5529 AAGCCGAGGCTTCTGCTTCTCAGCAGTGTCTGCGGAGCCCGGCTCCCTGTCC 5588
6015 CCTTCGTCGAGTCTTGGAGGGTATCACCCAGGAGCAGGCTCACTGACGCCCTCC 6074
5589 CCTTCGTCGAGTCTTGGAGGGTATCACCCAGGAGCAGGCTCACTGACGCCCTCC 5648
6075 TCCCCACAGGCCCGCTGATGCTCGGGAGCCCTGCGCCCGCATGGAGCTCTTCTCT 6134
5649 TCCCCACAGGCCCGCTGATGCTCGGGAGCCCTGCGCCCGCATGGAGCTCTTCTCT 5708
6135 TCTTCACTCTCTGCTGACGACTTACAGTCTTCTCGTGGCGCGCGACAGCCCGGCCCA 6194
5709 TCTTCACTCTCTGCTGACGACTTACAGTCTTCTCGGTCCTCACTGACAGCCCGGCCCA 5768
6195 GGCACCTGTGTGTCGTCAGTCTTCTGTCGACCCCATCCCTACAGGCTTGTGTCGTCG 6254
5769 GGCACCATGTGTGTCGTCAGTCTTCTGTCGACCCCATCCCTCATGAGCTTGTGTCGTCG 5828
6255 CCGCTAGAGTGGGTACCTAGTCTCCCGAGCCTGCTCCCTAGCCAGAGGCTCTTAATGTACA 6314
5829 CCGCTAGAGTGGGTACCTAGTCTCCCGAGCCTGCTCCCTAGCCAGAGGCTCTTAATGTACA 5888
6315 ATAAAGCAATGTGGTAGTTCACCTTGGGTCCCTGCTCAACGCTCTGTTGGATCATCC 6374
5889 ATAAAGCAATGTGGTAGTTCACCTGCGGTCCCTGCTCAACGCTCTGTTGGATCATCC 5948
6375 TCCCTAGGCGCAACCCACCCCTGCTCATCTGCTTACCCACCGCTGCGGCAATTTG 6434
5949 TCCCTAGGCGCAACCCACCCCTGCTCATCTGCTTACCCACCGCTGCGGCAATTTG 6008

6435 AGAC--GGGTAGCTTGGCTGAGCAGATGTGAGTTTACCTTGGCCCATATCCCATGTCC 6492
6009 AGACAGGGGTAGCTTGGCTGAGCAGATGTGAGTTTACCTTGGCCCATATCCCATGTCC 6068
6493 CCCACTGACCCCACTGACTGCCAGATTGGTGAAGAAGTACATTTGCTCGGATGT 6552
6069 CCCACTGACCCCACTGACTGCCAGATTGGTGAAGAAGTACATTTGCTCGGATGT 6128
6553 GGGGAAGGGGGCAGAAATGGCTGACTAGAGGTGTGAGTCCAGCCCTGGATGTGGAGAG 6612
6129 GGGGAAGGGGGCAGAAATGGCTGACTAGAGGTGTGAGTCCAGCCCTGGATGTGGAGAG 6188
6613 GGCAGAGTCTGAGCTGAGAGGCCCAATATTTTCAGGCTTAACTCAGCCCAATCAGGG 6672
6189 GGCAGAGTCTGAGCTGAGAGGCCCAATATTTTCAGGCTTAACTCAGCCCAATCAGGG 6248
6673 ACAGCAGTCTGCGAGCACCATCAACAGTCAACCTCCCTTATATATGACACCCCAAAA 6732
6249 ACAGCAGTCTGCGAGCACCATCAACAGTCAACCTCCCTTATATATGACACCCCAAAA 6308
6733 TGGGAAGCAAAATCAT---GTGAGGAGCTATATGCGAGG---CTACCTCCAGGGCTC 6785
6309 CGGAAGCAAAATCATGCGCTCAGGAGCTATATGCGAGGCTACCTACCTCCAGGGCTC 6368
6786 AGTCGAGGTGCCAGAAATTTCTCTGGGAAGGCCCCAGGAAACCCAGAGCCGAGCCAC 6845
6369 AGTCGAGGTGCCAGAAATTTCTCTGGGAAGGCCCCAGGAAACCCAGAGCCGAGCCAC 6428
6846 CGCCCTCAGCCT 6857
6429 CACCCTCAGCCT 6440

RESULT 9

AX687028 LOCUS AX687028 9433 bp DNA linear PAT 31-MAR-2003
DEFINITION Sequence 2 from Patent EP1281755.
ACCESSION AX687028
VERSION AX687028.1 GI:29409532

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 Milos, P.M. and Webb, S.M.
AUTHORS Variants of the human cyp2d6 gene
TITLE Patent: EP 1281755-A 2 05-FEB-2003;
JOURNAL Pfizer Products Inc. (US)

FEATURES
Location/Qualifiers
1..9433
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 51.7%; Score 5314.6; DB 6; Length 9433;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 5895; Conservative 0; Mismatches 269; Indels 149; Gaps 27;

637 TGTACAAAAAATCAGCCAGGCTAGTGGTGGGACCTGTAATCCAGTACTTTGGAGGC 696
186 TGAATAACAAAAAGCTAGACGCTGGTGGCACAACCTGTAATCCAGTACTTTAGGAGGC 245

697 TGAGCAGAGAAATCGCTTGAACCTCGAGGTGGAGTTGCAGTGGAGCCGAGATCGATC 756
246 TGAGCAGAGAAATTCCTTGAAGCCCTAGAGGTGAAGGTTGTAGTGAAGCCGAGATTCATC 305

757 ATTGCA-----CTCCCACTGGGCGAC-AGAGCAAAATCTCGTCTC----- 796
306 ATTGCAAAATGAGGGGAGCCACAGCTGGGCAACAGAGGAATCTCCGCTCCAAAA 365

797 -----AAAAAAAATTTATTTGGGCTCGGTGGTGGTCTGTAGTCCAGCTACTTGGAG 849

366	AAAAA	AAAAAAAAAAAGAAATTAGGCTGGGTGGTCTAGTGCCAGCTACTTTGGAG	425
850	GCAGAGGTCCACTTGATGTTGAGATTGCAGTGCAGGCCATGATCTCGCCACTGCACATCCGG	909	
426	GCAGGGGTCCACTTGATGTCGAGCTGCAGTGAGCCATGATCTGCGCATGTCACATCCGG	485	
910	CCCCGGCAAACAGAGTGCAGACCCTCTCTAAAGAAAAATAAAAATAAAAAAGCAACATATC	969	
486	CCTGGGCAAACAGAGTGCAGACCCTGTCTTAAGAA-----AAAAAAATAAGCAACATATC	540	
970	CTAAATAAAGGATCCTCCATAAATGTTTCCAACGATTTCTTAATCAGAAAAATCGAGGCCA	1029	
541	CTGAACAAGGATCCTCCATAACGTTCCACAGATTTCTTAATCAGAAAAATCGAGGCCA	600	
1030	GGAGCAGTGGAGATACGACCTCTCAGSCAGCCCTGGAGGATGCTGTACAGGCTGGGG	1089	
601	GAAAGCAGTGGAGAGACGACCTCTCAGSCAGCCCGGGAGGATGTTGTCAACGGCTGGGG	660	
1090	CAAGGGCTTTCAAGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTTGCAAAACAGGAAGT	1149	
661	CAAGGGCTTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTTGCAAAACAGNAGC	720	
1150	CATGGCCCGCCAGAGCCCCA-GAATGTGGGCTGAGCTGGGATCATGTGCAGCTTTGAG	1208	
721	CATAGCCCGCCAGAGCCCAGGAATGTGGGCTGGGCT-----	757	
1209	GCTCACCGGAGCAGCCTCTGGACAGAGAGGTCCTCATCCAGNAACCTCGGCATGGCT	1268	
758	-----GGGAGCAGCCTCTGGACAGAGTGGTCCCATCCAGAAACCTCCGCGATGGCT	810	
1269	GGGAAGTGGGTACTTGGTCCGGGCTCTGATGTGTGTGACTGTGGTGTGTGTGAGAGAG	1328	
811	GGGAAGTGGGTACTTGGTCCGGGCTCTGATGTGTGTGACTGTGTGTGTGTGTGAGAGAG	870	
1329	AATGTGTCCCTGAGTGCAGTGTGAGTCTGTGTGTGTGTGTGAATATTGTTTGTGTGGG	1388	
871	AATGTGTGCCCTAAGTGTGAGTGTGAGTCTGTGTGTGTGTGAATATTGTTTGTGTGGG	930	
1389	TGATTTTTTCGATGTGTAAATCGTGTCCCTGCAAGTGTGAACAAGTGGCAAGTGTCTGGG	1448	
931	TGATTTTTTCGCTGTGTAAATCGTGTCCCTGCAAGTGTGAACAAGTGGCAAGTGTCTGGG	990	
1449	AGTGGACAAGATCTGTGCACCATCAGTGTGTGCATAGCGTCTGTGCATGTCAAGAGT	1508	
991	AGTGGACAAGAGATCTGTGCACCATCAGTGTGTGCATAGCGTCTGTGCATGTCAAGAGT	1050	
1509	GCAAGGTGAAGTGAAGGGAACAGCCCATGATGCCACTCATCATCAGGAGCTCTAAGGCC	1568	
1051	GCAAGGTGAAGTGAAGGGAACAGSCCATGATGCCACTCATCATCAGGAGCTCTAAGGCC	1110	
1569	CCAGGTAAGTGCAGTGCACAGATAAGGGTCTGAGSTCACTCTGAGTGGGAGAGTGGG	1628	
1111	CCAGGTAAGTGCAGTGCACAGATAAGGGTCTGAAAGGTCACTCTGAGTGGGAGAGTGGG	1170	
1629	GGTAGGAAAGGGCAAGGTCAATGTTCTGGAGGAGGGGTGTGTGACTACATPAGGGTGTATG	1688	
1171	GGTAGGAAAGGGCAAGGCCATGTTCTGGAGGAGGGGTGTGTGACTACATPAGGGTGTATG	1230	
1689	AGCCTAGCTGGGAGGTGGATGGCCGGGTCCACTGAGACCTTGTTATCCAGAAAGCCTGT	1748	
1231	AGCCTAGCTGGGAGGTGGATGGCCGGGTCCACTGAAACCTTGTTATCCAGAAAGCCTTT	1290	
1749	GTGGGCTTGGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCGACACAGGCTTTCTTAC	1808	
1291	GCAGGCTTCCAGAGCTTGGAGTGGGAGAGGGGGTGACTTCTCCGACAGGCCCCCTCCAC	1350	
1809	CACCTACCTGGGTAAAGGCTCTGGACAGGAAGACAGCGGCAAGGACCTCTGGAGCAGCC	1868	
1351	CGGCTTACCTGGGTAAAGGCTTGGACAGGAAGAGGGGGCAAGAACCCTCTGGAGCAGCC	1410	
1869	CATACCTGCGCTGAGCTCTGCGCACTGGCAGCAAGTCAACACAGCAGGTTCACTCA	1928	

Db	1411	CATACCCGCCCTGGCCTGACTCTGCCACTGCGACGACAGTCAAACAGCAGGTTCACTCA	1471
QY	1929	CAGCAGAGGCGAAGGCCATCATCAGCTCCCTTTATAAGGGAAGGTCACGCGCTCGGTG	1988
Db	1471	CAGCAGAGGCAAGGCCATCATCAGCTCCCTTTATAAGGGAAGGTCACGCGCTCGGTG	1530
QY	1989	TGCCGAGAGTGTCCTGCTCGTGGTCTCTGTGCTCGTGGGTGGGGTGGCCAGGTGTGTCC	2048
Db	1531	TGCTGAGAGTGTCTGCTCGTGGTCTCTGTGCTCGTGGGTGGGGTGGCCAGGTGTGTCC	1590
QY	2049	AGAGGAGCCAGTTGGTAGTGAGCGACCATGCGGGCTGAAGCACTGGTGTCCCTGGGCCA	2108
Db	1591	AGAGGAGCCATTTGGTAGTGAGCGAGTATGGGGCTAGAAAGCACTGGTGTCCCTGGCCG	1650
QY	2109	TGATAGTGGCCATCTTCCTGCTCTGCTGACCTGATGCAACCGGCAACAGCTGGGCTG	2168
Db	1651	TGATAGTGGCCATCTTCCTGCTCTGCTGACCTGATGCAACCGGCGCAACGCTGGGCTG	1710
QY	2169	CACGCTACCGCCAGGTCCCTGCGCATGCGCGGGCTGGGCAACCTTGTGTGATGTGGAC	2228
Db	1711	CACGCTACCCACGAGGCCCTCGCCACTGCGCGGCTGGGCAACC-TGCTGCAATGTGGAC	1769
QY	2229	TTCCAGAAACACCATACTGCTCTGACCAAGGTGAGGAGGAGGTCTCTGAGGGCGGAGAG	2288
Db	1770	TTCCAGAAACACCATACTGCTCTGACCAAGGTGAGGAGGAGGTCTCTGAGGGCGGAGAG	1829
QY	2289	GGTCTCAGAGATGCCCAACACAGCAAAACATCGGTGGTGTGGGTAAACACAGGCTGGAT	2348
Db	1830	GGTCTCAGGCTCCCTTACCAGAAAGCAACATGATGGTGGGTGAAACACAGGCTGGAC	1889
QY	2349	CAGAAGCCAGGCTGAGAAGGGGAAGCAGGTTTGGGGGACGTTCTCTGGGAAGGACATTTA	2408
Db	1890	CAGAAGCCAGGCTGAGAAGGGGAAGCAGGTTTGGGGGACG-TCTTGGGAAGGGCATTTA	1948
QY	2409	TACATGCGATGAAGGACTGGAATTTCCAAAGGCCAAGGAAGATAGGGCAAGGGCCTGGA	2468
Db	1949	TACATGCGATGAAGGACTGGAATTTCCAAAGGCCAAGGAAGATAGGGCAAGGGCCTGGA	2008
QY	2469	GGTGGAGCTGGACTTGGCAGTGGGCAATGCAAGGCCATTTGGGCAACATATGTTATGAGTA	2528
Db	2009	GGTGGAGCTGGACTTGGCAGTGGGCAATGCAAGGCCATTTGGGCAACATATGTTATGAGTA	2068
QY	2529	CAAAAGTCCCTTCTGCTGACACAGAAAGGAAGCCCTTTGGGAATGGGAAGATGAGTTAGTCC	2588
Db	2069	CAAAAGTCCCTTCTGCTGACACAGAGGAAGGCCCTTTGGGAATGGGAAGATGAGTTAGTCC	2128
QY	2589	TGAGTCCGCTTTAAATCAAGAAATCGAGGATGAAGGGGFTGAGTGACCCGGTTCAAACC	2648
Db	2129	TGAGTCCGCTTTAAATCACGAAATCGAGGATGAAGGGGFTGAGTGACCCGGTTCAAACC	2188
QY	2649	TTTTTGCACTGGGTCTCGGGCCTCACTG-CTCACCGGCATGGACATCATCTGGGAAT	2707
Db	2189	TTTTTGCACTGGGTCTCGGGCCTCACTGCTCACCGGCATGGACATCATCTGGGAAT	2248
QY	2708	GGGATGCTAACTGGGGCCTCTCGGCAATTTTGGTGAATTTTGGCAAGTCACTCGGGTG	2767
Db	2249	GGGATGCTAACTGGGGCCTCTCGGCAATTTTGGTGAATTTTGGCAAGTCACTCGGGTG	2308
QY	2768	ACGATCCAAACTGAGTTCTCCATCAAGAAAGGTGACCCCAACCCCTGCCCCACGAT	2827
Db	2309	ACGATCCAAACTGAGTTCTCCATCACAGAAAGGTGACCCCAACCCCTGCCCCACGAT	2368
QY	2828	CAGGAGGCTGGGTCTCTCTCTCCACTGCTCACTCTGTAGTGGCGGGGTGCTCCAA	2887
Db	2369	CAGGAGGCTGGGTCTCTCTCTCACTGCTCACTCTGTAGTGGCGGGGTGCTCCAA	2428
QY	2888	GGTTCAAATAGGACTAGGAACCTGTAGTCTGGGGTGAATCTTGGCTTGACAAGAGGCCCTGA	2947
Db	2429	GGTTCAAATAGGACTAGGACTGTAGTCTGGGGTGAATCTTGGCTTGACAAGAGGCCCTGA	2488
QY	2948	CCCTCCCTCTGCAAGTTGCGGCGCGCTTGGGGAACGTGTTAGCCTTGCAAGCTGGCCTGGA	3007
Db	2489	CCCTCCCTCTGCAAGTTGCGGCGCGCTTGGGGAACGTGTTAGCCTTGCAAGCTGGCCTGGA	2548

Qy	2525	AGTACAAAGTCCCTTCTGCTGACACACAGAGGAAAGCCCTTGGGAATGGAAGATGAGTTA	2584
Db	1259	AGTACAAAGTCCCTTCTGCTGACACACAGAGGAAAGCCCTTGGGAATGGAAGATGAGTTA	1318
Qy	2585	GTCTGAGTGCCGTTTAAATCAGAAATCGAGGATGAAGGGGTGCAGTGACCCGGTTC	2644
Db	1319	GTCTGAGTGCCGTTTAAATCAGAAATCGAGGATGAAGGGGTGCAGTGACCCGGTTC	1378
Qy	2645	AACCTTTTGCACTGTGGGTCTCTGGGCCTCACTGCTCACCGGATGAGACCATCATCTGGG	2704
Db	1379	AACCTTTTGCACTGTGGGTCTCTGGGCCTCACTGCTCACCGGATGAGACCATCATCTGGG	1438
Qy	2705	AATGGGATGCTAACTGGGGCCTCTCGGCAATTTTGGTGACTCTTTGCAAGGTCAATCTGG	2764
Db	1439	AATGGGATGCTAACTGGGGCCTCTCGGCAATTTTGGTGACTCTTTGCAAGGTCAATCTGG	1498
Qy	2765	GTGACGATCAAACTGAGTTCTTCATCAACAGAGGTGTGA	2824
Db	1499	GTGACGATCAAACTGAGTTCTTCATCAACAGAGGTGTGA	1558
Qy	2825	GATCAGAGGCTGGGTCTCTCTTTCACCTGCTCACTCTGCTAGCCCCGGGGGTGCTC	2884
Db	1559	GATCAGAGGCTGGGTCTCTCTTTCACCTGCTCACTCTGCTAGCCCCGGGGGTGCTC	1618
Qy	2885	CAAGGTTCAAAATAGACTAGGACCTGTAGTCTTGGGGTGATCCTGGCTTGA	2944
Db	1619	CAAGGTTCAAAATAGACTAGGACCTGTAGTCTTGGGGTGATCCTGGCTTGA	1678
Qy	2945	TGACCTTCCCTCTGCACTGGGGCCGCTTTCGGGGACGTGTTCAGGCTGCAGCTGGCCT	3004
Db	1679	TGACCTTCCCTCTGCACTGGGGCCGCTTTCGGGGACGTGTTCAGGCTGCAGCTGGCCT	1738
Qy	3005	GGACGCGGTGGTCTGCTCAATGGGCTTGGCGCGTGGCGGAGGCGATGGTCAACCGCG	3064
Db	1739	GGACGCGGTGGTCTGCTCAATGGGCTTGGCGCGTGGCGGAGGCGATGGTCAACCGCG	1798
Qy	3065	GCGAGGACACGGCCCGCTCGCGCCCATCTACAGAGTCTGGGCTTTCGGGGCCG	3124
Db	1799	GCGAGGACACGGCCCGCTCGCGCCCATCTACAGAGTCTGGGCTTTCGGGGCCG	1858
Qy	3125	GTTCCCAAGGACAGCGCGGTGGGGGACAGAGACCGGTTTCCGTGGGCCCCCGGTGGAC	3184
Db	1859	GTTCCCAAGGACAGCGGTGGGGGACAGAGACAGAGAGAG-ATTTTCGTGGGACCCCGGTGGGT	1916
Qy	3185	AGTCACCGTAGCCCAACAGCGCCGACAGAGCGGTGGGGTCTTGACGTGTAACACAGAGATA	3244
Db	1917	GATGACCGTAGTCCGAGCTGGGACAGAGGCGCGGGGTCTGTAACATGAACACAGGCCAG	1976
Qy	3245	AAGCCACGAGTGGGTGAGGACAGTGGGCGAGGAAACCACTGACGCGGGGAGGTGCG	3304
Db	1977	CGAGTGGGACAGCGGGCCAGAAACCACTGACATGAGGAGGTGTGAGCATGGGACGA	2036
Qy	3305	AGTCTGTG--GGCTGGAGGGGGCGGGCTTACTGCGCCAGACCCGCGCAGAAAGCCCGTGGG	3362
Db	2037	GGGCGGGCTGTGACAGTGGGCGGGGCCACTGCGAGACCTGGCAGGAGGCCAATGGG	2096
Qy	3363	CGAGGCTATGC-----GTGCAAGTGGCGGTGGCGGGGACCCGGCC	3403
Db	2097	TGAGCTGGGCGCATTTCCAGCTGGAATCCGGTGTGCAAGTGGGGGCGGGGACCCGACCC	2156
Qy	3404	TATGCTCGGGCTCAGTGTGG-----CGGGGACGGGCGGATCTTCTTCAGTGGGAAAGG	3458
Db	2157	TGTGCTGTAGCTCAGTGTGGGTGGCGGGGGCCCGGGGGTCTTCCCTGAGTGCAGAGG	2216
Qy	3459	TGGTGAGGTTGGGACAGAGAGTGGGGCCAAACCCCGCCCAAGGACAGGGGACCAATGT	3518
Db	2217	CGGTGAGGTTGGGACAGAGACGAGGT--GGGCAAGCCCTGCGCCACAGCGGACGCAAGGT	2275
Qy	3519	GGGTGACCAAGACTGGGCCCTGTGCCAGCTGGAACCGGGCTAGGACCTCGGGGAGACCT	3578
Db	2276	GGATGCAAAAGAGTGGGCCCTGTGACACAGCTGGACAGAGCCAGGGACTCGGGGAGACCA	2335
Qy	3579	TGTGGAGCGCCAGGGTTTGGAGTGGGTGGCGGAGGGGTGGGGCCAAAGGCTTTCATGGCAACG	3638

2336	Db	GGGGGAGC-ATAGGGTTGGAGTGGGTGGTGGATGGTGGGGCTAATGCTTTCATGGCCAGC	2394
3639	Qy	CCACAGTGTCCGTCGCGCCGCCACAGGGGTGATCTGTCTGGCTATGGCCCGCGTGGCGCG	3698
2395	Db	CGCAGTGGCCGTCCTCACCCCAAGGGTGTCTCTGGCGCTATGGCCCGCGTGGCGCG	2454
3699	Qy	AGCAGAGGCGCTTCTCCGTGTCCACCTTGGCGAACTTGGGCTTGGGCAAGAACTCGCTGG	3758
2455	Db	AGCAGAGGCGCTTCTCCGTCTCACCTTGGCGAACTTGGGCTTGGGCAAGAACTCGCTGG	2514
3759	Qy	AGCAGTGGGTGAACGAGAGGCGCGCTTCTTGTGCGG-CTTTCGCGCAACAAAGCCGGTG	3817
2515	Db	AGCAGTGGGTGAACGAGAGGCGCGCTTCTTGTGCGCGCTTCTTGTGCGCGCAACCACTCGG	2574
3818	Qy	GGTGATGGGCAGAAAGGCAAAAGCGGAACTGGGAAGCGGGGACGGGAAGGCAAC	3877
2575	Db	GGTGATGGGCAGAAAGGCAAAAGCGGAACTGGGAAGCGGGGACGGGAAGGCAAC	2634
3878	Qy	CTTTACCGCATCTCCCAACCCACAGACGCCCTTTTCGCCCAAGCGGCTCTTGACAAC	3937
2635	Db	CTTTACCGCATCT-CCCAACCCCAAGACGCCCTTTTCGCCCAAGCGGCTCTTGACAAC	2693
3938	Qy	AGCCGTGAGCAACGTGATGCCCTCCCTCACTCGCGGCGCGCTTCAGAGTACGACACCC	3997
2694	Db	AGCCGTGAGCAACGTGATGCCCTCCCTCACTCGCGGCGCGCTTCAGAGTACGACACCC	2753
3998	Qy	TGCGTTCTCAGAGCTGTGACCTAGCTCAGGAGGACCTGAGGAGAGTTCGGGCTTCT	4057
2754	Db	TGCGTTCTCAGAGCTGTGACCTAGCTCAGGAGGACCTGAGGAGAGTTCGGGCTTCT	2813
4058	Qy	GGCGAGGTGCGAGCGAGAGACCGCAGGAGTCTCTGACGGGCGAGCTCTCTGAGAGTGCC	4117
2814	Db	GGCGAGGTGCGAGCGAGAGACCGCAGGAGTCTCTGACGGGCGAGCTCTCTGAGAGTGCC	2873
4118	Qy	GGGCGTGAACCTGGGGCTTCGAAAGGCGAGATTTGCATAGATCGGTTTGGGAAAGACAT	4177
2874	Db	GGGCGTGAACCTGGGGCTTCGAAAGGCGAGATTTGCATAGATCGGTTTGGGAAAGACAT	2933
4178	Qy	TCCAGGAGACCCACACTGTAAAGAGGCGCTGGAGAGGGGAGACATCTCAGACATGGTCG	4237
2934	Db	TCCAGGAGACCCACACTGTAAAGAGGCGCTGGAGAGGGGAGACATCTCAGACATGGTCG	2993
4238	Qy	TGGAGAGGTGTGCCCGGGTCAGGGGCGACACGAGAGGCGCAAGAGCTCTGTACC-CCCC	4296
2994	Db	TGGAGAGGTGTGCCCGGGTCAGGGGCGACACGAGAGGCGCAAGAGCTCTGTACCCTTA	3053
4297	Qy	TCCAAGTGTGAGATTTGATTTAGTTTCTCCTCTGGGCAAGGAGAGGGTGGAGGCT	4356
3054	Db	TCCAAGTGTGAGATTTGATTTAGTTTCTCCTCTGGGCAAGGAGAGGGTGGAGGCT	3113
4357	Qy	GGCATTGGGGAGGACCTTGGTGAAGTCAAGTGAAGGACAGGAGGCGCTTGGTCTACC	4416
3114	Db	GGCATTGGGGAGGACCTTGGTGAAGTCAAGTGAAGGACAGGAGGCGCTTGGTCTACC	3173
4417	Qy	TGAGATGGCTGGGCGCTGAGACTGTGTCCAGGTGAACGAGACACAGGAGGGATTTGAGA	4476
3174	Db	TGAGATGGCTGGGCGCTGAGACTGTGTCCAGGTGAACGAGACACAGGAGGGATTTGAGA	3233
4477	Qy	CCCCGTTCTGTCTGGTGTAGGTGCTGAATGCTGTCCCCGCTCCTCTGCACATCCCCAGCGC	4536
3234	Db	CCCCGTTCTGTCTGGTGTAGGTGCTGAATGCTGTCCCCGCTCCTCTGCATATCCCCAGCGC	3293
4537	Qy	TGGTGGCAAGGTCTTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGAGCTGCTAA	4596
3294	Db	TGGTGGCAAGGTCTTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGAGCTGCTAA	3353
4597	Qy	CTGAGCACAGATGACCTGGGACCCACGCCAGACCCCGGAGACCTGACTGAGGCGCTTCC	4656
3354	Db	CTGAGCACAGATGACCTGGGACCCAGGCCAGGCCCTCCGAGACCTGACTGAGGCGCTTCC	3413
4657	Qy	TGGCAAGAAAGGAGAGTGAGAGTGGCTGCCACGTTGGGGGCAAGGGTGGTGGGTGA	4716

Db 3414 TGGCAGAGATGGGAAGGTGAGTGGCTGCCACGGTGGGGGGCAAGGGTGGTGGTTGA 3473
Qy 4717 ACCTCCAGAGGAATGAGGGAGGCTGGGCAAAAGTTGGACAGTGCATCACCCGGGG 4776
Db 3474 GCGTCCAGAGGAATGAGGGAGGCTGGGCAAAAGTTGGACAGTGCATCACCCGGGG 3533
Qy 4777 AGCCGATCTGGGCTGACAGGTGACAAATGGAGGTCAATTTGGGGGCTACCCGTTCTAT 4836
Db 3534 AGCCGATCTGGGCTGACAGGTGACAAATGGAGGTCAATTTGGGGGCTACCCGTTCTGT 3593
Qy 4837 CCCCTCAGTATCTCTCGGCCCTCTCAGGCCAAGGGAGCCCTGAGAGCAGCTCAATG 4896
Db 3594 CCC - -GAGTATGCTTCGGCCCTCTCAGGCCAAGGGAGCCCTGAGAGCAGCTCAATG 3651
Qy 4897 ATGAGAACCTGCGCATAGTGGTGGGTAACTGTTCTTGCCTGGGATGGTGACACCTCGA 4956
Db 3652 ATGAGAACCTGCGCATAGTGGTGGGTAACTGTTCTTGCCTGGGATGGTGACACCTCGA 3711
Qy 4957 CACGCTGGGCTGGGCTCTCTGCTCATGATCTTACCTAGATGTGACGCTGAGCCCA 5016
Db 3712 CACGCTGGGCTGGGCTCTCTGCTCATGATCTTACCTAGATGTGACGCTGAGCCCA 3771
Qy 5017 GCTGGGGCCCAAGGCAAGGACTGAGGAGGAGAGGATACAGCTGGGGGCCCTGGGCTTAG 5076
Db 3772 TCTGGGAACAAGTGCAGGGGCCGAGGGAGAGGGTACAGGGGGGGCCCATGAACCTTTG 3831
Qy 5077 CTGGGACACCGGGCTTCCAGCACAGGGGTGCGCAGGCTCTCTGTAAGCTTAACCTTCCTC 5136
Db 3832 CTGGGACACCGGGCTTCCAGCACAGGCTTACACAGGATCTGTAAAGCTTACCTCTCTC 3891
Qy 5137 CAAACAGAGAGAGAGAGTGT-----CCCCCTGGGTGCTGACCCATTTGTGG 5183
Db 3892 CAAACATAGAGGCAAGAGAGGTGTACAGGGCCGGAACCCCTGGGTGCTGACCCATTTGTGG 3951
Qy 5184 GAGCGATCTCTGCAGTCCGTTCCAAACAGAGATGCAAGCGTGAAGGCAAGGCTGC 5243
Db 3952 GAGCGATCTCTGCAGGCGGTGTCACAGAGAGATCAGACGATGATAGGGCAGGTGC 4011
Qy 5244 GCGACACAGAGATGGGTGACAGGCTCACATGCCCTTACACCACTGCGGTGATTACAGAG 5303
Db 4012 GCGACACAGAGATGGGTGACAGGCTCACATGCCCTTACACCACTGCGGTGATTACAGAG 4071
Qy 5304 TGAGCGCTTTGGGACATCATCCCCCTGAGTGTGACCATATGACATCCCCGTGACATCG 5363
Db 4072 TGAGCGCTTTGGGACATCATCCCCCTGAGTGTGACCATATGACATCCCCGTGACATCG 4131
Qy 5364 AAGTACAGGCTTCCGATCCCTAAGTGGCTGGGCGCTCTCACCAGCTCAGCA 5423
Db 4132 AAGTACAGGCTTCCGATCCCTAAGTGGCTGGGCGCTCTCACCAGCTCAGCA 4191
Qy 5424 CCAGCACTGGTGATAGCCCAAGCATGGCTACTGCGAGGTGGGCCCACTTAGGAACCT 5483
Db 4192 CCAGCCCTGGTGATAGCCCAAGCATGGCTACTGCGAGGTGGGCCCACTTAGGAACCT 4251
Qy 5484 GGCACCTAGTCTCAATGCCACCACTGACTGTCCCCCTTGGGTGGGGGTCCAGAG 5543
Db 4252 GGCACCTAGTCTCAATGCCACCACTGACTGTCCCCCTTGGGTGGGGGTCCAGAG 4311
Qy 5544 TATAGCAGGCTGGGCTGCTCCATCAGAGCCCCCTCTAGTGGGGAGACAAACACAGAC 5603
Db 4312 TATAGCAGGCTGGGCTGCTCCATCAGAGCCCCCTCTAGTGGGGAGACAAACACAGAC 4371
Qy 5604 CTGCGAAGATGTTGGAGGACCAAGCCCTGAGGGAGAGGGGAGGGGAGGCTGCTCTG 5663
Db 4372 CTGCGAAGATGTTGGAGGACCAAGCCCTGAGGGAGAGGGGAGGGGAGGCTGCTCTG 4431
Qy 5664 AGAGGTGTGACTGCGCCCTCTGCTGGGGTGGAGAGGGTACTGTGGAGCTTCTCGGGGCG 5723
Db 4432 AGAGGTGTGACTGCGCCCTCTGCTGGGGTGGAGAGGGTACTGTGGAGCTTCTCGGGGCG 4491
Qy 5724 AGGACTAGTTGACAGAGTCCAGCTGTGCGCAGGAGTGTGTCTCCCGCTGTTTGT 5783
Db 4492 AGGACTAGTTGACAGAGTCCAGCTGTGCGCAGGAGTGTGTCTCCCGCTGTTTGT 4551

Qy 5784 GCGAGGGTCCAGCATCTAGAGTCCAGTCCCACTCTCACCTGCATCTCTGCCCCAG 5843
Db 4552 GCGAGGGTCCAGCATCTAGAGTCCAGTCCCACTCTCACCTGCATCTCTGCCCCAG 4611
Qy 5844 GGAACGACATCTATCACCAACCTGTCTGCTGAGGATGAGGCGCTCTGGGAGAG 5903
Db 4612 GGAACGACATCTATCACCAACCTGTCTGCTGAGGATGAGGCGCTCTGGGAGAG 4671
Qy 5904 CCCTTCGCTTCCACCCCGAACACTTCTTGAATGCCAGGCGCACTTCTGAAAGCCGGAG 5963
Db 4672 CCCTTCGCTTCCACCCCGAACACTTCTTGAATGCCAGGCGCACTTCTGAAAGCCGGAG 4731
Qy 5964 GCGTTCGCTTCTTCAGCAGGTGCTGTGGGGAGCCCGGCTCCCTGTCCCCCTTCCGTG 6023
Db 4732 GCGTTCGCTTCTTCAGCAGGTGCTGTGGGGAGCCCGGCTCCCTGTCCCCCTTCCGTG 4791
Qy 6024 GAGTCTTGAGGGGTATCACCCAGAGCCAGGCTCACTACAGCCCTCTCCCTCCCCACAG 6083
Db 4792 GAGTCTTGAGGGGTATCACCCAGAGCCAGGCTCACTACAGCCCTCTCCCTCCCCACAG 4851
Qy 6084 GCGCGCTGTCATGCTCGGGAGCCCTGCGCCCGCATGAGGCTCTTCTCTTCTTCCACT 6143
Db 4852 GCGCGCTGTCATGCTCGGGAGCCCTGCGCCCGCATGAGGCTCTTCTCTTCTTCCACT 4911
Qy 6144 CCTCTCTGAGCATCTTTCAGCTTCTCGGTGGCCCGCGGACAGCCCGGCGCCAGCACTCTC 6203
Db 4912 CCTCTCTGAGCATCTTTCAGCTTCTCGGTGGCCCGGACAGCCCGGCGCCAGCACTCTC 4971
Qy 6204 GTGCTGTCAGCTTCTGCTGACCCCATCCCTCTAGAGCTTGTGCTGTGCCCGCTAGA 6263
Db 4972 GTGCTTCTGCTTCTGCTGACCCCATCCCTCTAGAGCTTGTGCTGTGCCCGCTAGA 5031
Qy 6264 ATGGGTACTAGTCTCCAGCCTGCTCCCTAGCAGAGCTCTAATGTAACAATAAAGCAA 6323
Db 5032 ATGGGTACTAGTCTCCAGCCTGCTCCCTAGCAGAGCTCTAATGTAACAATAAAGCAA 5091
Qy 6324 TGTGTAGTGTCCAACTTGGGTCCCTGCTCAGCCCTCTGTTGGGATCATCTCTCAGGG 6383
Db 5092 TGTGTAGTGTCCAACTTGGGTCCCTGCTCAGCCCTCTGTTGGGATCATCTCTCAGGG 5151
Qy 6384 CAAACCCACCTGCTCATCTGCTTACCCCAAGCCCTGCGCGGATTTGAGAC--GGG 6441
Db 5152 CAAACCCACCTGCTCATCTGCTTACCCCAAGCCCTGCGCGGATTTGAGACAGGGG 5211
Qy 6442 TAGGTTGAGGCTGAGCAGATGTACCTTACCTTGCCTAATCCCATGTCCCCCACTGAC 6501
Db 5212 TAGTGTGAGGCTGAGCAGATGTACCTTGCCTAATCCCATGTCCCCCACTGAC 5271
Qy 6502 CCAACTCTGACTGCCCAGATTGGTGACAGGACTACATTGTCTGGCATGTGGGAGGG 6561
Db 5272 CCAACTCTGACTGCCCAGATTGGTGACAGGACTACATTGTCTGGCATGTGGGAGGG 5331
Qy 6562 GCCAAGATGGGTGACTAGAGGTGACTGACCCCTGGATGTGGTGGAGAGGCGAGACT 6621
Db 5332 GCCAAGATGGGTGACTAGAGGTGACTGACCCCTGGATGTGGTGGAGAGGCGAGACT 5391
Qy 6622 CAGCCCTGGAGGCCCATATTTAGGCGCTTAACCTCAGCCCAAGCCCACTCAGGAGCAGCTC 6681
Db 5392 CAGCCCTGGAGGCCCATATTTAGGCGCTTAACCTCAGCCCAAGCCCACTCAGGAGCAGCTC 5451
Qy 6682 CTGCCAGCACCATCAACAGTCACTCCCTTTCATATATGACACCCCAAAA 6732
Db 5452 CTGCCAGCACCATCAACAGTCACTCCCTTTCATATATGACACCCCAAAA 5502

RESULT 11
AX345458
LOCUS
DEFINITION Sequence 529 from Patent WO0200928.
ACCESSION AX345458
VERSION AX345458.1 GI:18493344
KEYWORDS

AX345458 5884 bp DNA linear PAT 01-FEB-2002

```
SOURCE      synthetic construct
ORGANISM     synthetic construct
              artificial sequences.
REFERENCE    1
AUTHORS      Olek,A., Piepenbrock,C. and Berlin,K.
TITLE        Diagnosis of diseases associated with the immune system
JOURNAL      Patent: WO 0200928-A 529 03-JAN-2002;
              Epigenomics AG (DE)
FEATURES     Location/Qualifiers
              source
              1..5884
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"
              /note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN
Query Match      18.8%; Score 1927.6; DB 6; Length 5884;
Best Local Similarity 78.4%; Pred. No. 0;
Matches 2322; Conservative 0; Mismatches 639; Indels 1; Gaps 1;

QY      1  AGCCTTACAAGTGTGGGATACCTCGTGTAGCCACCGGGTCCGGCTCTTTATGCTTT 60
Db      2924  AGTTTTATAAGAGTGTGGGATTAATTTGGTGAATATCGGGTTCGGTTTTTTTATGTTTT 2983

QY      61  ACTGTACTGTCTGTCTTGAAGAGTACTTATTTTGGATTGATTCATCATTTAGCTAA 120
Db      2984  ATTGATTTGTTGTTTGAAGAGTATTTATTTATTTTGGATTGATTCATCATTTAGCTAA 3043

QY      121  TTAATAAAGAGTAGTTTACACACCAATTAAGTATTATTAATCTCTGTTTTCTGTG 180
Db      3044  TTAATAAAGAGTAGTTTATATATTAATTAATTAATTAATTAATTAATTTGTTTTG 3103

QY      181  TCGTACTATTACAGTAGTGTGTTGACCTTTAGATGATTCCTCTGCTCATTAATATC 240
Db      3104  TCGTTATTATTATAGTAGTGTGTTGTTTATTTAGATGATTTTGTGTTTATTAATAT 3163

QY      241  CTTTTTTTTTTCAGATTGAAGAACTCCCTTTAGCATTTCTGTGGATATAGCTGTG 300
Db      3164  TTTTTTTTTTGTAGATTGAAGAAATTTTTTTTGTAGTATTTTTTGTGGATATAGGTTG 3223

QY      301  TTGATGAATCTCGCAGCTTTTGTGTTGCTGGGAAGGCTTTATTTCTCTCTCTCTG 360
Db      3224  TTGATGAATCTCGTAGTTTTTGTGTTGTTGGAAGGTTTTTATTTTTTTTGTG 3283

QY      361  AAGGATATTTTCCAGATACGTTATCTAGGCTAAAGTTTTTTTTTCTCCCTCAGCATT 420
Db      3284  AAGGATATTTTGTAGATACGTTATTTTGTAGTTAAAGTTTTTTTTTTTATTTT 3343

QY      421  AAATATGTCATCCACTCCCTCCCTGCTGAAGGTTTCCACTGGAAGGTCGTCGCC 480
Db      3344  AAATATGTTATGTTATTTTTTTTGTGTTTGTGAAGGTTTTATTTGAAGGTTGTTT 3403

QY      481  ATGTCATGTAATGGAGCTCTACTGCATGTTATTTGTTCTTTCTCTGCTCTTTAGG 540
Db      3404  ATGTTATGTAATGGAGTTTATTTGATGTTATTTGTTTCTTTTGTGTTGTTTAGG 3463

QY      541  ATCCCTTTCTTTATCTTGACCTTCGGAGTTTAATTAICAGATGCCCTGAGTCTCTTC 600
Db      3464  ATTTTTTTTTTATTTTGAATTTTTCGGAGTTTAATTAATAGATGTTTTCAGGTCGTTT 3523

QY      601  TTTGGGTTAAATCTGCTGCTGCTCTATAAATCTCTGTGACAAAATCAGCCAGGCATG 660
Db      3524  TTTGGGTTAAATGTTTGGTGTTTATTAATTTTTTTGTTATTAATAAATTAATGTTAG 3583

QY      661  GTGTGGGCACTGTAATCCAGCTACTTGGAGGCTGAGGAGGAGTAATCGCTTTGAACC 720
Db      3584  GTGTGGGATTTTGTAAATTTTATGTTATTTGGAGGTTGAGGTAGGAGTAATCGTTGA 3643

QY      721  CTGGAGGTGGAGGTTGCAGTAGCCGAGATCCCATCATTTGCACTCCACCTGGGGCAG 780
Db      3644  TTGGAGGTGGAGGTTGTAGTAGTCGAGATCGTATTTATTTTATTTATTTATTTT 3703

QY      781  AGCAAAACTCCGCTCTCAAAAAAAATTTATTTGGGCTCGGTGCTGCTAGTCCAGCT 840

3704  AGTAAATTTTGGTTTTAAAAAAATTTATTTGGTTCGGTTCGTTGTTAGTTTAGTT 3763
841  ACTTGGGAGGAGGAGTCCACTTGATGTTGAGATTGAGTGGAGCATGATCCTCCACT 900
3764  ATTTGGGAGGAGTGGAGGTTTATTTGATGTTGAGATTGAGTGGAGTTATGATTTGTTAT 3823
901  GCATCCCGGCGGGCAACAGAGTCAGACCCTGTCTAAAGAAAAATTAATAATAAAG 960
3824  GTATTTTCGGTTTCGGGTAAATAGAGTGAGATTGTTGTTTAAAGAAAAATTAATAAAG 3883
961  CAACATATCTTAAATAAAGGATCTCCATAATGTTTCCACAGATTTCTTAATCAGAAACA 1020
3884  TAATATATTTAAATAAAGGATTTTATATATGTTTATTAATGATTTTATTAATAAATA 3943
1021  TGGAGGCCAGGAAGCAGTGGAGAAATGACGACCCCTCAGCAGCCCTGGAGGATGCTGTAC 1080
3944  TGGAGGTTAGGAAGTAGTGGAGAAATGACGATTTTATAGGTAGTTTGGAGGATGTTGTTAT 4003
1081  AGGCTGGGGCAAGGGCTTTCAGGCTACCACTGGGAGCTCTGGGACAGCCCTGTTGCA 1140
4004  AGGTTGGGGTAAGGGTTTTTATAGGTTTAAATTTGGAGGTTTTTGGGAATAGTTTGTGTA 4063
1141  ACAGGAAGTCTATGGCCCGGCGCAGAGCCAGAAATGAGGCTGAGCTGGGATCCATGTGACA 1200
4064  ATAGGAAGTATGTTTCTGTTAGAGTTTAGAATGTTGGGTTGAGTTGGGATTTATGTGATA 4123
1201  GCTTTGAGGCTCACCGGAGCAGCCTCTGGACAGGAGGTCCTCCATCAGGAACCTCGG 1260
4124  GTTTTGAAGTTATCGGGAGTAGTTTTTGGATAGGAGAGGTTTTTATTTAGGAATTCGG 4183
1261  GCATGGCTGGGAAGTGGGGTACTTGGTCCGGGCTCTGATGTTGTTGTCAGCTGGTGTG 1320
4184  GTATGGTTGGGAAGTGGGGTATTTGGTGTGCGGTTTTGATGTTGTTGATGTTGGTGTG 4243
1321  TGAGAGAAATGTTGTCCTCAGTGTCTAGTGTGAGTCTGTGATGTTGTTGAAATTTGCTT 1380
4244  TGAGAGAAATGTTGTTTGGAGTTAGTGTGAGTTGTTGTTGTTGTTGTTGTTGTTT 4303
1381  TGTGTGGTGTATTTTCTCATGTGTAATGCTGTCTCCCTCAAGTGTGAACAAGTGACAA 1440
4304  TGTGTGGGTGATTTTTTGTATGTTGTAATCGTGTGTTTTTCTAAGTGTGAATTAAGTGA 4363
1441  TGTCTGGAGTGGACAGAGATCTGTGCACCATCAGGTTGTTGTCATAGGCTGTGTCATG 1500
4364  TGTGTGGAGTGGATTAAGAGATTTGTTGTTATTTATAGGTTGTTGTTATAGGTTTGTG 4423
1501  TCAAGAGTGAAGGTGAAGTGAAGGACCGGCCCATGCTCCACTCATCATCAGGAGCT 1560
4424  TTAAGAGTGAAGGTGAAGTGAAGGATTAAGGTTTATGATGTTATTTATTTATTTAGAG 4483
1561  CTAAGGCCCCAGGTAAGTGCAGTGACAGATAAGGGTCTGAAAGTCTCTGAGAGTGGG 1620
4484  TTAAGGTTTTAGGTAAGTGTAGTATAGATAAGGGTCTGAAAGGTTTATTTGGAGTGGG 4543
1621  CAGGTGGGGTAGGGAAGGCAAGGTCATGTTCTGGAGGAGGGGTTCTGACTACATTAG 1680
4544  TAGTGGGGGTAGGGAAGGTTAAGGTTATGTTTGGAGGAGGGGTTGTTGATTATTTAG 4603
1681  GGTGTATCAGCCTAGCTGGGAGGTGGAATGGCCGGCTCCACTCAGACCCCTGGTTATCC 1740
4604  GGTGTATCAGTTTAGTTGGGAGGTGGAATGGCTGGGTTTATTCAGATTTTGGTTATTT 4663
1741  AAGCCTGTGTGGCTTGGGGAGCTTTGGAGTGGGGAGAGGGGTGACTCTCCGACAGGC 1800
4664  AAGTTGTGTGGGTTTGGGAGGTTTGGAGTGGGAGAGGGGTGATTTTTTTCGATTAGT 4723
1801  CTTTCTACCACTTACCCTGGGTAAGGCTTGGAGCAGGAGCAGGAGGACCTCTG 1860
4724  TTTTATTTATTTATTTTGGTAAAGGTTTGGAGTAGGAGTAGGAGTAAAGGATTTT 4783
1861  GAGCAGCCCATACCTGCGCTCGCTGACTCTGCCACTGGCAGCAGTCAACACAGCAGG 1920
```

```
Db 4784 GAGTAGTTATATATTTGTTTGGTTTGAATTTTGTATTTGGTAGTATAGTTAATATAGTAGG 4843
Qy 1921 TTCACTCACAGCAGCGGCGAAGCCCATCATCAGCTCCCTTTTATAAGGGAAGGTCACGC 1980
Db 4844 TTTTATTTATAGTAGAGGCGAAGGTTATTTATTTAGTTTATTTTATTAAGGGAAGGTTACGC 4903
Qy 1981 GCTCGGTGTCGCGAGAGTGTCCTGCGCTGCTCTGTGCTGCTGTGGGTGGGGTGGCCAG 2040
Db 4904 GTTCGGTGTGTCGAGAGTGTTTTGTTTGGTTTTTGTGTTTGGTGGGTGGGGTGTGTAG 4963
Qy 2041 GTGTGTCAGAGCGCCAGTGTGTAGTAGGAGCAGCCATGCGGCTAGAACACTGTGTGCC 2100
Db 4964 GTGTGTTTAGAGAGTTAGTTGCTAGTAGGAGTTATTTGCGGTTAGAGTATTTGTTGT 5023
Qy 2101 CTTGGCCATGATAGTGGCCATCTCTCTGCTCTCTGCTGACCTGATGACCGGACCAACG 2160
Db 5024 TTTGGTTATGATAGTGGTTATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5083
Qy 2161 CTGGGCTGACGCTACCGCCAGGTCCCTGCCACTGCCCGGCTGGGCAACCTTGCTGC 2220
Db 5084 TTGGGTTGTACGTTATTTGTTAGGTTTTTTTGTATTTGTTGCGGTTGCGTAAATTTGTTGT 5143
Qy 2221 ATGTGACTTCCAGAACACACACTACTGCTTCGACGAGGTGAGGAGGAGTCTCGAGG 2280
Db 5144 ATGTGATTTTGAATAATATATATTTGTTTTCGATTTAGGTGAGGAGGAGGTTTTGGAGG 5203
Qy 2281 GCGCAGAGTCTCTGAGGATGCCCCACACACCAACATGCGTGGTGGGTTAAACACACA 2340
Db 5204 GCGGTAGAGGTTTGAGGATGTTTTTATTTAGTAAATATGCGTGGTGGGTTAAATATA 5263
Qy 2341 GCGTGAATCAGAACCGAGCTGAGAGGGGAGCAGGTTTGGGGGACGTTCTCTGGGGAAG 2400
Db 5264 GGTTCGATTTAGAAAGTTAGGTTGAGAAAGGGGAGTAGGTTTGGGGGACGTT-TTGGGGAAG 5322
Qy 2401 GACATTTATACATGGCATGAAGGACTGGATTTTCCAAAGCCAGGAAGTAGGCGCAG 2460
Db 5323 GATATTTATATATGTTAGAGGATTTGATTTTTTAAAGGTTAAGGAAGTAGGAGTAAG 5382
Qy 2461 GGCCTGGAGGTGAGCTGGACTTGGCAGTGGGATGCAAGCCCATTTGGGCAACATATGTT 2520
Db 5383 GGTTCGAGGTGAGTTGGATTTGGTAGTGGGTATGTAAGTTTATTTGGGTATATATGTT 5442
Qy 2521 ATGGAGTACAAATCCCTCTCTGTGACACAGAAAGAAAGGCTTGGGAATGGAAGATGA 2580
Db 5443 ATGGAGTATAAAGTTTTTTTGTGTGATATTAGAAAGAAAGGTTTGGGAATGGAAGATGA 5502
Qy 2581 GTTAGTCTCAGTGCCTTTAAATACGAAATCGAGGATGAAGGGGTGAGTGACCCCG 2640
Db 5503 GTTAGTTTTGAGTGTCTTTAAATACGAAATCGAGGATGAAGGGGTGAGTGATTCGG 5562
Qy 2641 TTCAAACTTTTGCACTGTGGGTCTCTCGGCTCCTGCTCACCGGCATGGACCATCATC 2700
Db 5563 TTTAAATTTTGTATTTGGGTTTTTGGGTTTTATTTGTTATCGGTATGGATATATTT 5622
Qy 2701 TGGGAATGGGATGCTAACTGGGCTCTCGGCAATTTTGGTGACTCTTGCAGAGTCAATAC 2760
Db 5623 TGGGAATGGGATGTTAAATGGGGTTTTTCGGTAAATTTTGGTGATTTTGTAGGTTATAT 5682
Qy 2761 CTGGGTGAGCATCCAAACTGAGTTCCTCATCAGAAAGGTGTGACCCCGCCCTGCC 2820
Db 5683 TTGGGTGAGCGTATTTAAATTTGAGTTTTTTTATATAGAAAGGTGTGATTTTTTTTCGTT 5742
Qy 2821 CCACGATCAGGAGGTGGGTCTCTCTCTCCACTGCTCAGTCCCTGAGTACCCCGGGGT 2880
Db 5743 TTAGGATTAGGAGGTGGGTTTTTTTTTTTATTTTATTTTGGTAGTTTCGGGGGT 5802
Qy 2881 CGTCCAAGGTTCAAAATAGGATAGGACCTGTAGTCTGGGGTGTATCTCGCTTGACAAAGAG 2940
Db 5803 CGTTTAAAGTTTAAATAGGATTAGGATTTGTAGTTTGGGGTGTATTTGTTTGTATAGAG 5862
Qy 2941 GCGCTGACCCCTCCTCTGAGT 2962
Db 5863 GTTTTGAATTTTTTTTTTTGTAGT 5884
```

```
RESULT 12
AX348344
LOCUS AX348344 5884 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 39 from Patent WO0202806.
ACCESSION AX348344
VERSION AX348344.1 GI:18614380
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE
  1
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Method and nucleic acids for pharmacogenomic methylation analysis
JOURNAL Patent: WO 0202806-A 39 10-JAN-2002;
FEATURES
    Epigenomics AG (DE)
    Location/Qualifiers
      source
      1..5884
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN
Query Match 18.8%; Score 1927.6; DB 6; Length 5884;
Best Local Similarity 78.4%; Pred. No. 0;
Matches 2322; Conservative 0; Mismatches 639; Indels 1; Gaps 1;
Qy 1 AGCCTTACAAAGTGTGGGATTTACCTGCGTGAGCCACCGGTCGGCCTCTTTATGCTTT 60
Db 2924 AGTTTATTAAGAGTTGGGATTTATTTGCGTGAGTTATCGGTTCCGTTTTTTATGTTT 2983
Qy 61 ACTGTACTGCTGCTGGAAGAGTACTTATTTTGGTTGTTGTTGTTGTTGTTGTTGTTGTT 120
Db 2984 ATTGATTTGTTGTTTGAAGAGTATTTATTTTGTGATTTGTTGTTGTTTATTTAGTTAA 3043
Qy 121 TTAAATTAAGAGTAGTTTACACACCAATTAAGTATTAATATACTCTGTTTCTGTTG 180
Db 3044 TTAAATTAAGAGTAGTTTATATAATTAATTAATTAATTAATTAATTAATTAATTAAT 3103
Qy 181 TCGTACTAATTAACAGTAGTGTGTTGACCTTTAGATGATTTCTTCTGCTCATATAATC 240
Db 3104 TGTATTATTAATAGTAGTGTGTTGTTTATTTTAGATGATTTTGTGTTGTTTATTAATATT 3163
Qy 241 CTTTCTTTTCTCAGATGGAAGTCTCCCTTTAGCATTTCTTGTGGATATAGTCTGGTG 300
Db 3164 TTTTCTTTTCTCAGATGGAAGTCTCCCTTTAGCATTTCTTGTGGATATAGTCTGGTG 3223
Qy 301 TTGATGAATCTCGCAGCTTTTGTGTTGCTGGGAAGGTCCTTTATTTCTCTCTCTGTTGG 360
Db 3224 TTGATGAATCTCGCAGCTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3283
Qy 361 AAGGATATTTTGTCCAGATACGTTATTTAGGCTAAAAGTTTTTTTCTTCAGCACTTT 420
Db 3284 AAGGATATTTTGTGTAGATACGTTATTTAGGCTAAAAGTTTTTTTCTTCAGCACTTT 3343
Qy 421 AAATATGATCGCACTCCCTCCCTGCTGAGGTTTCCACTGGAAGGTGCTGCCCC 480
Db 3344 AAATATGATCGCACTCCCTCCCTGCTGAGGTTTCCACTGGAAGGTGCTGCCCC 3403
Qy 481 ATGTCATGTTTGGAGCTCTACTGATGTTATTTGTTTCTTTCTCTCTCTCTCTCTCTCTCT 540
Db 3404 ATGTCATGTTTGGAGTTTATTTATGTTATTTGTTTCTTTCTCTCTCTCTCTCTCTCTCT 3463
Qy 541 ATCTCTTCTTTATCTCTGACCTTTCCGAGTTTAAATATATCAGATGCCCTGAGGTCTCTTC 600
Db 3464 ATCTCTTCTTTATCTCTGACCTTTCCGAGTTTAAATATATCAGATGCCCTGAGGTCTCTTC 3523
Qy 601 TTTGGTTAAATCTGCTTGGTGTCTATTAACCTCTTGTACAAAATCAGCCAGGCATG 660
Db 3524 TTTGGTTAAATCTGCTTGGTGTCTATTAACCTCTTGTACAAAATCAGCCAGGCATG 3583
```


Db 5743 TTAGGATTAGGAGTTGGGTTTTTTTTTTTATTTATTTTGTATTTTGTAGTTTCGGGGT 5802

Qy 2881 CGTCCAAGTTCAAAATAGGACTAGGACCTCTAGTCTGGGGTGATCTCGGCTTGACAAGAG 2940

Db 5803 CGTTTAAGGTTAAATAGGATTAGGATTTCTAGTTTGGGGTGATTTTGGTTTGAAGAG 5862

Qy 2941 GCCTGACCCCTCCCTCTGCACT 2962

Db 5863 GTTTTGATTTTTTTTTTTTGTAGT 5884

RESULT 13

AX345459/c

LOCUS AX345459 5884 bp DNA linear PAT 01-FEB-2002

DEFINITION Sequence 530 from Patent WO0200928.

ACCESSION AX345459

VERSION AX345459.1 GI:18493345

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS Olek A., Piepenbrock C. and Berlin K.

TITLE Diagnosis of diseases associated with the immune system

JOURNAL Patent: WO 0200928-A 530 03-JAN-2002;

Epigenomics AG (DE)

FEATURES

Location/Qualifiers

source 1..5884

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN

Query Match 15.9%; Score 1631.6; DB 6; Length 5884;

Best Local Similarity 72.1%; Pred. No. 0;

Matches 2137; Conservative 0; Mismatches 824; Indels 1; Gaps 1;

Qy 1 AGCCTTACAAAGTCTGGGATTACCTGCGTGAGCCACCGGTCGGGCTCTTTATCTCTT 60

Db 2961 AACCTTACAANAATCTAANAATCTAGTACGTAACACCGGATCCGACCTCTTTATCTT 2902

Qy '61 ACTGTACTGTCTGTGAAAGTACTTATTTTGTGATGTTTCTATCATTTAGTCTAA 120

Db 2901 ACTATACTATCTTAAATAATCTTATTTTAAATTAATTCATCATTTAACTTAA 2842

Qy 121 TTAANAATAGAGTAGTTTACACCAACCAATACAGTATTAATATCTGTTTCTG 180

Db 2841 TTAANAATAAANAATTAATTTACACCAACCAATTAACAATATTAATATCTTATTTTCTATA 2782

Qy 181 TGCTTACTATTACAGTAGTTTGTGTACCTTTAGATGATTTCTTTGTCTCATTAAATATC 240

Db 2781 TACTTACTATTACCAATAAATTTTATACCTTTTAAATTAATTTCTTTTACTCATTAAATC 2722

Qy 241 CTTTCTTTTTCAGATTGAAAACTCCCTTTAGCATTTCTTTGGGATATAGTCTGGTG 300

Db 2721 CTTTCTTTTTCAGATTGAAAACTCCCTTTAGCATTTCTTTTAAATTAATTAATCTAATA 2662

Qy 301 TTGATGAATCTCGAGCTTTTGTCTGCGGAAGTCTTTATTTCTCTTCTCTGTTGG 360

Db 2661 TTAATAAATCTCGCACTTTTATTTATCTTAAATAAATCTTTATTTCTCTCTCTTCTATTA 2602

Qy 361 AAGGATATTTTCCAGATAGCTTATCTAGGCTTAAAGTTTTTTTCTTCTAGCACTTT 420

Db 2601 AAAAATATTTTACCATAACGTTATCTTAAACTTAAATAAATTTTCTTCTTCAACACTTT 2542

Qy 421 AAATATGTCATGCACTCCCTCCCTGCTGTAAGTTTCCACTGGAAGTGCTGCCCC 480

Db 2541 AAATATATCATACACTCCCTCCCTTAACCTTAAATAAATTTTCCACTTAAATAAATTAACCTCC 2482

Qy 481 ATGTCATGTATGGAGCTCTACTGCAATGTTATTTGTTTCTTTCTCTGCTGTTTAGG 540

Db 2481 ATATCATATTAATAAATCTACTACATATTTATTTATTTTCTTTCTCTTACTACTTTTAA 2422

Qy 541 ATCTTTTCTTTATCTCTTGACCTTTTGGAGTTTAAATTTATCAGATGCTTGAGTCTGCTTC 600

Db 2421 ATCTTTTCTTTATCTCTTAACTTTTGAATTTAAATTTATCAATACCTTAAATTCGCTTC 2362

Qy 601 TTTGGTTAAATCTGCTTGGTCTTATAAATCTTCTTGTACAAAAATCAGCCAGGCTG 660

Db 2361 TTTAAATTAATCTACTTAAATTTCTTATAAATCTTCTTATAAATAATCAACAAACATA 2302

Qy 661 GTGGTGGCACCTGTAAATCCAGCTACTTTGGGAGGCTGAGGAGGAGAAATCGGTGTAACC 720

Db 2301 ATAATAAACACCTTAAATCTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2242

Qy 721 CTGGAGGTGGAGTTTGCAGTGCAGCCGAGATCGCATCATTTGCATCTCCACCTGGGCGACAG 780

Db 2241 CTAAAAATTAATAATTAACAATAAACCGAAATCGCATCATTTACACTCCCACTTAAACGACA 2182

Qy 781 AGCAAAACTCCGTCTCAAAAAAATAATTTTGGGCTCGGTGGTCTGTAGTCCGAGCT 840

Db 2181 AACAAAACTCCGTCTCAAAAAAATAATTTTAAACTCGATATACTCTATAATCTCCAACT 2122

Qy 841 ACTTGGAGGAGGAGGTCCACTTGTATTTGAGATTGAGTTCAGTGCAGTGCAGTGCATCTG 900

Db 2121 ACTTAAAAAACAATAAATCCACTTATAATTTAAATTTACATAAACCATAATCTCTACACT 2062

Qy 901 GCACCTCCGCGCGGCAACAGAGTGCAGCTGTCTTAAAGAAAAATAAATAAATAAATAA 960

Db 2061 ACACCTCCGCGCGGCAACCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2002

Qy 961 CAACATATCTTAAATAAAGATCTCTCATTAATTTTCCACAGATTTCTTAATCAGAAACA 1020

Db 2001 CAACATATCTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1942

Qy 1021 TGGGCGCAGGAGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGCAGGAGTGCAG 1080

Db 1941 TAAAAACCAAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1882

Qy 1081 AGGCTGGGCGGAGGCGCTTTCAGGCTACCAACTGGGAGCTCTGGGAAACAGCCCTGTTGCAA 1140

Db 1881 AAATTAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1822

Qy 1141 ACAGGAAGTCTATGCGCGCGGCGGAGAGCCAGAGATGTGGGTGAGTGGGATGCCATGTGACA 1200

Db 1821 ACAAAAAATCATAAACCCGACCAAAACCCAAATAATAAACTAAACTAAATAAATAAATAAATAA 1762

Qy 1201 GCTTTGAGGCTCACCGGAGGAGCTCTGAGCAGGAGGAGTCCCATCCAGGAAACCTCGG 1260

Db 1761 ACTTTAAAACTCACCGGAAACCAACCTCTTAAACAAAAAATAAATAAATAAATAAATAAATAA 1702

Qy 1261 GCATGGCTGGGAAGTGGGTAATTTGGTGGCGGCTCTGTATGTGTGTGACTGGTGTGTG 1320

Db 1701 ACATAACTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1642

Qy 1321 TGAGAGAGATGTGCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1380

Db 1641 TAAAAAATAATATATACCTTAAATCAATAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1582

Qy 1381 TGTGTGGGTGATTTTCTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTG 1440

Db 1581 TATATAAATAAATTTTCTACATATAATCTGATCTCCCTACAAATAAATAAATAAATAAATAAATAA 1522

Qy 1441 TGTCTGGGAGTGCAGAGAGATCTGTGCACCATCTGAGTGTGTCATAGGCTGTGTCATG 1500

Db 1521 TATCTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1462

Qy 1501 TCAAGAGTCAAGGTTGAAGTGAAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCT 1560

Db 1461 TCAAAAAATCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1402

Qy 1561 CTAAGGCCCAAGTGAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1620

Db 1401 CTAAAAACCCCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1342

```
QY 1621 CAGGTGGGGTGGAGGAGGCAAGGTGCTGTTCTGGAGGAGGGTGTGACTACATTAG 1680
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1341 CAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 1282
QY 1681 GGTGTATGAGCCTAGCTGGGAGGTGGATGCGCGGTCCACTGAGACCCCTGTTATCCCAAG 1740
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1281 AATATATAAACCTTAACATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 1222
QY 1741 AAGCCTGTGTGGCTTGGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCGACAGGC 1800
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1221 AAACCTATATAAATCTTAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATA 1162
QY 1801 CTTTCTACCACTCCCTGGGCTAGGGCTGGAGCAGGAAGCAGCGGCAAGGACTCTG 1860
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1161 CTTTCTACCACTCCCTTAATAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 1102
QY 1861 GAGCAGCCATACCTGCCCTGCTGACTCTGCCACTGSCAGCAGACTCAACACAGCAGG 1920
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1101 AAACAACCCATACCTACCTAACTCTACCACTAACACACAATCAACACACAA 1042
QY 1921 TTCCTACAGCAGAGGGGGAAGGCATCATCAGCTCCCTTTTATAGGAAGGCTCAGC 1980
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1041 TTCACTCACAACAAAACCAACCAATCAATCAATCAATCAATCAATCAATCAATCAAT 982
QY 1981 GCTCGGTGCGCAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
981 GCTCGATATACCGAAATAATCTTACTTAATCTCTATACCTTAATAAATAAATAAATA 922
QY 2041 GTGTGTCCAGAGGAGCCAGTGTGTAGTGAGGAGCAGCCAGTGGGCTAGAGCACTGGTGC 2100
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
921 ATATATCCAAAACCAACCAATTAATAAATAAACAACCAATAAATAAATAAATAAATA 862
QY 2101 CTTGGCATGATAGTGGCCATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
861 CCTAACCAATAAATAAACCATCTTCTACTCTAATAAACCTTAATAAATAAATAAATA 802
QY 2161 CTGGGTGACAGCTACCCGCCAGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
801 CTAACCTACAGCTACCCGCCAATTCCTCTACCTACCTACCTACCTACCTACCTACCT 742
QY 2221 ATGTGGACTTCAGAACACACATATCTGCTTCCAGCAGGTGAGGAGGAGTCTGGAGG 2280
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
741 ATATAAATCTCCAAAACACACATATCTTCCAGCAATAAATAAATAAATAAATAAATA 682
QY 2281 GCGCAGAGGTCTGAGGATGCCCCACAGACAACTGCTGGTGGTGGTGGTGGTGGTGGTGG 2340
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
681 ACGACAAAATCTTAAAAATACCCCAACCAACCAATAAATAAATAAATAAATAAATAA 622
QY 2341 GCGTGATCAGAGCCAGGCTCAGAGGGGAGCAGGTTTGGGGGACGTTCTCTGGGAG 2400
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
621 AACTAAATCAAAACCAAACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 563
QY 2401 GACATTTATACATGCGATGAAGGACTGGATTTTCCAAAGGCCAAGGAGAGTAGGCGAAG 2460
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
562 ACATTTATACATACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 503
QY 2461 GCGCTGGAGGTGGAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2520
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
502 AACCTAAAAATAAACTAACTTAACAATAAATAAATAAATAAATAAATAAATAAATAA 443
QY 2521 ATGGAGTACAAAGTCCCTTCTGCTGACACCAAGGAAGGCTTGGGAATGGAAGATGA 2580
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
442 ATAAATACAAATCCCTTCTACTAACCAAAATAAATAAATAAATAAATAAATAAATAA 383
QY 2581 GTTAGTCTGAGTGGCTTTAAATCAAGAAATCGAGGATGAAGGGGTGAGTGGAGCCGG 2640
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
382 ATTAATCTTAATACCGTTTAATCAAGAAATCGAAATGAATAAATAAATAAATAAATA 323
QY 2641 TTCAAAACCTTTTGCACTGTTGGTCTCTGGGCTCTACTGCTCAACCGCATGGAACATATC 2700
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
322 TTCAAAACCTTTTACATATAATCTCTGAACCTCACTACTCAACCGCATGAACATATC 263
QY 2701 TGGGAATGGATGCTAACTGGGGCCTCTCGGCAATTTTGTGACTCTTGGCAGGTCATAC 2760
```

```
Db 262 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 203
QY 2761 CTGGGTGAGCATCCAAACTGAGTCTCCATCAGAGAGGTGTGACCCCAACCCCTGCGC 2820
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
202 CTAATAAAGCATCCAAATAAATCTCTCCATCAGAAAAATAAATAAATAAATAAATAAATA 143
QY 2821 CCACGATCAGAGGCTGGGTCTCTCTCTCCACCTGCTCAGTCTCTGGTAGCCCGGGGT 2880
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
142 CCAAAATCAAAAAATAAATCTCTCTCTCCACCTACTCCTCTTAATAAACCCTGAAAT 83
QY 2881 CGTCCAAAGTTCAAATAGGACTAGGACCTAGTCTGGGGTGCATCTCGCTGTGACAAGAG 2940
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
82 CGTCCAAATTCAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 23
QY 2941 GCCCTGACCCCTCCCTCTGAGT 2962
Db 22 ACCCTAACCCCTCCCTCTACAAT 1

RESULT 14
AX348345/c
LOCUS
DEFINITION
Sequence 40 from Patent WO0202806.
ACCESSION
AX348345
VERSION
AX348345.1 GI:18614381
KEYWORDS
synthetic construct
synthetic construct
artificial sequences.
SOURCE
ORGANISM
Olek, A., Piepenbrock, C. and Berlin, K.
AUTHORS
Method and nucleic acids for pharmacogenomic methylation analysis
TITLE
Patent: WO 0202806-A 40 10-JAN-2002;
JOURNAL
Epigenomics AG (DE)
FEATURES
Location/Qualifiers
1..5884
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN
Query Match 15.9%; Score 1631.6; DB 6; Length 5884;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 2137; Conservative 0; Mismatches 824; Indels 1; Gaps 1;

QY 1 AGCCTTACAAAGTGTGGGATTACTCGCTGAGCCAGCCGGTCCGGCTCTTTATGTCTT 60
Db 2961 AACCTTACAAATACTAAATAATCTAGTGTAAACCCGAAATCCGACCTCTTTATATCTT 2902
QY 61 ACTGTACTGTCTGTCTGAAAAAGTACTTATTATTTTGTGATTTTCATCATTTAGTCTAA 120
Db 2901 ACTACTACTATCTATCTTAAAAAATACTATTATTATTTTAAATTAATTCATCATTTAATCTAA 2842
QY 121 TTAATAATGAAGTAGTGTTTACACACCAATTTACAGTATATATAATCTCTGTTTCTGTG 180
Db 2841 TTAATAATAAAAATAATTTACACCAATTTACAAATATATAATACTCTATTTTTCTATA 2782
QY 181 TGCTTACTATTACCAAGTGTGTTTGTACTTTAGATGATTTCTTCTGCTCATTAATATC 240
Db 2781 TACTTACTATTACCAATATAATTTTATACCTTTAAATAATTTCTTCTACTCATTAATATC 2722
QY 241 CTTTCTTTTTCAGATTGAAAAATCCCTTTAGCATTTTCTGTGGATATAGGTCGTGGTG 300
Db 2721 CTTTCTTTTTCAAAATTAATAAACTCCCTTTTAAACATTTCTTATAAAATAAATAAATAAATA 2662
QY 301 TTGATGAATCTCGCAGCTTTTGTGCTCGGAAGGTCTTTTATTTCTTCTCTCTGTTGG 360
Db 2661 TTAATAAAATCTCGCACTTTTATTATCTAAAAAATACTTTTATTTCTCTCTCTTAAATAA 2602
QY 361 AAGGATATTTTGGCAGATAGTATTCTAGGCTAAAAAGTTTCTTCTCTCAGCACTTT 420
```


Db 1319 CGCATAGTGGTGGCTGACCTGTCTCTGCCGGATGGTGACCACTCGACCAACGCTGGCC 1378
 Qy 4968 TGGGGCTCCTGCTCATGATCCTACACCTGGATGTGCAGCGTGAGCCAGCTGGGGCCCA 5027
 Db 1379 TGGGGCTCCTGCTCATGATCCTACATCCGGATGTGCAGCGTGAGCCATCTGGGAAACA 1438
 Qy 5028 AGCAGGGACTG 5039
 Db 1439 GTCAGGGGCCG 1450

Search completed: February 28, 2004, 11:39:56
 Job time : 25617.9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 25, 2004, 02:12:03 ; Search time 432.346 Seconds
(without alignments)
13433.798 Million cell updates/sec

Title: US-09-820-788A-3
Perfect score: 18731
Sequence: 1 agccttaaaagtctggga.....ccagggtcagtcgaggt 10278

Scoring table:
BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/USPTO spool_p/US09820788/runat_24022004_141425_20017/app_query.fasta_1.12174
-DB=A_Geneseq_25Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09820788 @CGN 1.1 859 @runat_24022004_141425_20017 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_29Jan04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1807	9.6	502	6 ABU09595	Abu09595 Human cyt
2	1803	9.6	502	6 ABU09594	Abu09594 Human cyt
3	1802	9.6	497	2 AAR72376	Aar72376 Human aux
4	1802	9.6	497	2 AAR93183	Aar93183 Human cyt
5	1800	9.6	497	6 ADB25832	Adb25832 Human CYP
6	1799	9.6	497	2 AAW44869	Aaw44869 Cytochrom
7	1799	9.6	497	6 ABU09598	Abu09598 Human cyt
8	1798	9.6	497	2 AAR72377	Aar72377 Human aux
9	1798	9.6	497	2 AAR93184	Aar93184 Human cyt
10	1798	9.6	497	2 AAR81462	Aar81462 Human der

11	1798	9.6	497	5	AB009563	Human CYP
12	1798	9.6	497	6	AA026405	Human CYP
13	1798	9.6	497	6	ADB25833	Human CYP
14	1795	9.6	497	6	ABU09593	Human cyt
15	1795	9.6	497	6	ABR82026	Human cyt
16	1795	9.6	497	7	ADG60558	Human Pro
17	1794	9.6	497	2	AAR72375	Human aux
18	1794	9.6	497	2	AAR93182	Human cyt
19	1790	9.6	497	2	AAR72378	Human aux
20	1790	9.6	497	2	AAR93185	Human cyt
21	1790	9.6	497	6	ADB25834	Human CYP
22	1786	9.5	497	6	ADB25831	Human CYP
23	1682	9.0	497	4	AAE05171	Human dru
24	1551.5	8.3	446	6	AAO26404	Human dru
25	1316.5	7.0	436	4	AAU68579	Human nov
26	1273.5	6.8	500	7	ADE60556	Rat Prote
27	1126	6.0	504	6	AAO22644	Protein o
28	851	4.5	195	2	AAW44870	Cytochrom
29	823	4.4	184	4	AAG74361	Human col
30	740	4.0	139	4	AAO00527	Human pol
31	712.5	3.8	19938	6	ABP76679	Streptomy
32	686.5	3.7	326	4	ABG28250	Novel hum
33	666.5	3.6	159	2	AAW44871	Cytochrom
34	653.5	3.5	19938	6	ABP76679	Streptomy
35	627.5	3.4	19938	6	ABP76681	Streptomy
36	615.5	3.3	19938	6	ABP76678	Streptomy
37	614	3.3	330	4	AAO03028	Human pol
38	614	3.3	656	4	ABG27623	Novel hum
39	606	3.2	19938	6	ABP76682	Streptomy
40	601.5	3.2	19938	6	ABP76682	Streptomy
41	601	3.2	19938	6	ABP76680	Streptomy
42	599	3.2	390	4	ABG24889	Novel hum
43	598	3.2	587	4	ABG27019	Novel hum
44	598	3.2	682	4	ABG10590	Novel hum
45	594	3.2	19938	6	AB898398	Streptomy

ALIGNMENTS

RESULT 1
ABU09595
ID ABU09595 standard; protein; 502 AA.
XX AC ABU09595;
XX AC

16-JUL-2003 (first entry)

Human cytochrome p450 gene CYP2D6, variant G5799C/C5816AT, protein.

Human; cytochrome P450; CYP2D6; chromosome 22; SNP;
single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
psychiatric disorder; drug sensitivity.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 486

FT /note= "Wild-type Ser substituted by Thr, as the result of a single nucleotide polymorphism"

FT Misc-difference 492..502

FT /note= "These amino acids replace the wild-type sequence (Leu-Cys-Ala-Val-Pro-Arg) as the result of a single-nucleotide polymorphism which causes a frameshift"

XX EP1281755-A2.

PN 05-FEB-2003.

PD 16-JUL-2002; 2002EP-00254972.

PR 31-JUL-2001; 2001US-0309111P.

PA (PFIZ) PFIZER PROD INC.

XX MILOS PM, Webb SM;
XX WPI; 2003-373769/36.
DR N-PSDB; ACA61305.
XX
PT New cytochrome P450 2D6 gene variants and polypeptides, useful for
PT determining if a subject has or is at risk of developing a drug
PT sensitivity condition or disorder that is associated with an aberrant
PT CYP2D6 activity.
XX
PS Claim 4; Fig 8; 88pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising a cytochrome
CC P450 2D6 gene variant, e.g. G5799C or C5816A (referring to the genomic
CC sequence or the same variant nucleotide in the corresponding cDNA
CC sequences). Also included are probes, primers (allele specific
CC oligonucleotides) and arrays used to detect and/or amplify the CYP2D6
CC gene polymorphic regions, the variant polypeptides, antibodies which are
CC capable of distinguishing between the variant and wild-type polypeptides,
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
CC determining whether an individual is susceptible to being a poor
CC metaboliser of drugs. The DNA probe is useful for hybridising to a
CC variant form of the CYP2D6 gene. The primer is useful for amplifying the
CC C5816A allelic variant. The allele specific nucleotide is useful for the
CC detection of the C5816A allelic variant. The methods are useful for
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and
CC determining if an individual is susceptible to being a poor metaboliser
CC of drugs. The nucleic acids are useful as probes or primers for
CC determining whether a subject has a genetic deficiency for metabolising
CC drugs that are substrates of P450 CYP2D6. The methods are useful for
CC determining if a subject has or is at risk of developing a drug
CC sensitivity condition or disorder that is associated with an aberrant
CC CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
CC aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
CC appropriate drugs or determining the course of treatment to administer to
CC a subject to treat cardiovascular or psychiatric disorders, or for
CC treating a subject with a drug sensitivity or disorder associated with a
CC specific allelic variant of a polymorphic region of the CYP2D6 gene. The
CC antibodies are useful for monitoring CYP2D6 protein levels in an
CC individual for determining whether a subject has a disease or conditions
CC associated with an aberrant CYP2D6 protein level. The gene is located on
CC human chromosome 22. The present sequence is the G5799C/C5816A double
CC variant CYP2D6 protein
XX
SQ Sequence 502 AA;

Alignment Scores:
Pred. No.: 502 Length: 502
Score: 1807.00 Matches: 469
Percent Similarity: 34.31% Conservative: 13
Best Local Similarity: 33.38% Mismatches: 20
Query Match: 9.65% Indels: 904
DB: 6 Gaps: 9

US-09-820-788A-3 (1-10278) x ABU09595 (1-502)

QY 2078 ATGGGGCTAGAGCACTGGTGGCCCTGGCCATGATAGTGGCCATCTTCTGCTCTGCTG 2137
DB 1 MetGlyLeuGluAlaLeuValProLeuAlaValAlaValAlaLeuLeuVal 20
QY 2138 GACCTGATGACCGGACCAACGCTGGCTGACGCTACCGCTACCGCCAGGTCCTCGCACTG 2197
DB 21 AspLeuMetHisArgGlnArgTrpAlaAlaGlyProProGlyProLeuProLeu 40
QY 2198 CCGGGCTGGCAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2257
DB 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspG1 60
QY 2258 GGTGAGGAGGAGGTCTTGGAGGCGGACAGAGGTCTTGGAGGATGCCCCACCAACGCAAA 2317

Db 60 n----- 60
QY 2318 CATGGTGGTGGGTTAAACACACAGCTGGATCAGAACCCAGGCTGAGNAGGGAGCAGG 2377
Db 60 ----- 60
QY 2378 TTTGGGGACGTTCTCTGGGGAAGGACATTTATACATGSCATGAAGGACTGGATTTTCCAA 2437
Db 60 ----- 60
QY 2438 AGGCCAAGGAAGAGTAGGAGGCAAGGGCCCTGGAGGTGGAGCTGGAGCTGGAGTGGGCGATGC 2497
Db 60 ----- 60
QY 2498 AAGCCATTGGGCAACATATCTTATGGAGTACAAAGTCCCTTCTGCTGACACGAGAAGGA 2557
Db 60 ----- 60
QY 2558 AAGGCTTTGGGAATGGAAGATGAGTTAGTCTCTGAGTCCGTTTAAATCAGAAATCGAGG 2617
Db 60 ----- 60
QY 2618 ATGAAGGGGTGTCAGTGCACCCGGTTCAAACTTTTGACACTGGGTCTCTCGGCGCTCACT 2677
Db 60 ----- 60
QY 2678 GCTCACCAGGATGAGCAATCATCTGGGAATGGGATGTAATCGGGGCTCTCGGCAATTT 2737
Db 60 ----- 60
QY 2738 TGGTGACTCTTTCAGAGTTCATACCTGGGTGACGATCCAACTGAGTTCTCTCCATCAG 2797
Db 60 ----- 60
QY 2798 AAGGTGTGACCCCAACCCCTGCCCAAGATCAGGAGGTGGGTCTCTCTTCCACCTGC 2857
Db 60 ----- 60
QY 2858 TCACCTCTGGTAGCCCGGGGGTGTCTCAAGTTCAATAGACTAGGACCTGTAGTCTG 2917
Db 60 ----- 60
QY 2918 GGGTGATCTGCTTGGTACAGAGGCCCTGACCCCTCTCTGAGTTGGGCGCGCTTCG 2977
Db 61 ----- -Leu-ArgArgArgPheG 66
QY 2978 GGGAGTGTTCAGCTGCTGAGTGGCTGGACGCGGTGCTGCTGCTCAATGGGCTGGCGG 3037
Db 66 lyAspValPheSerLeuGlnLeuAlaTrpThrProValValValLeuAsnGlyLeuAla 86
QY 3038 CCGTGGCGAGCGATGGTGAACCCGGCGGAGGACAGCGCGGCGGCGGCTGGCGCCA 3097
Db 86 laValArgGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProValProI 106
QY 3098 TCTACAGTCTCTGGGCTTCTGGGCGGCTTCCCAAGGCAAGCGGCGGTGGGGGACAGAGA 3157
Db 106 leThrGlnIleLeuGlyPheGlyProArgSer-Gln----- 117
QY 3158 CCGGCTTTCTCGTGGGCGGCGGCTGGACAGTGCACCGTAGCCCAAGCAGCGCGGCGG 3217
Db 117 ----- 117
QY 3218 TGGGGTCTTGGACGTGAACAGAGATTAAGCCAGCGAGTGGGTGAGGACAGTGGGCGCA 3277
Db 117 ----- 117
QY 3278 GGAACACCTGACGCGGGAGGTGCGAGTCTGTGGGCTGGAGGGGCGGGCTACTGC 3337
Db 117 ----- 117
QY 3338 CCAGACCCGCAAGAGCCCGGTGGGCGAGGCTGATGCTCGAAGTGGCGGTGGCGGGGAC 3397
Db 117 ----- 117

QY	3398	CGCGCCCTATGCTGCGGGCTCAGTGTGGGCGGACGGCGGGATCTTCTTTGAGTGAAG	3457	QY	4478	CCGTTCTCTGCTGAGTGTGCTGAATGCTGTCTCCCGTCTCTCTGACATCCAGCGCT	4537
Db	117	-----	117	Db	223	-----ValLeuAsnAlaValProValLeuHisileProAlaLe	236
QY	3458	GTGGTCAGGGTGGCAGAGAGAGTGGGGCCAAACCCCGCCAGCAGGGGAGCAATG	3517	QY	4538	GGCTGCAAGTCTACGCTTCCAAAGGCTTTCCTGAGCCAGCTGAGTGGCTGAAC	4597
Db	117	-----	117	Db	236	uAlaGlyIysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh	256
QY	3518	TGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGGCTAGGGAATGCGGGAGACC	3577	QY	4598	TGAGCAGAGGATGACCTGGGACCCAGCCAGCCAGCCAGCCAGCTGAGTGGCTTCCT	4657
Db	117	-----	117	Db	256	rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe	276
QY	3578	TTGTGGAGCCAGGGTTGAGTGGGTGGCGAGGGTGGGCCAAGGCCCTTCATGGCAAC	3637	QY	4658	GGCAAGAGAGAGAGTGGCTGCTGCCAGCTGGGTGGGGGGCAAGGGTGGTGGGTGA	4717
Db	117	-----	117	Db	276	uAlaGluMetGluLys-----	281
QY	3638	GCCCACTGTCTCGTCCCGCCCGAGGGGTGATCTGTCTGCGCTATGGGCGCGTGGCGC	3697	QY	4718	CGTCCAGAGGAGAAAGAGGGAGGCTGGGCAAAAGTTGGACCAAGTTCATCAGCCCGCGA	4777
Db	118	-----GlyValPheLeuAlaArgTyrGlyProAlaTrpArg	129	Db	281	-----	281
QY	3698	GAGCAGAGGGCTTCTCCGTGTCCACCTTCGCAACTTGGGCTGGGCAAGAAAGTGCCTG	3757	QY	4778	GCCGATCTGGGCTGACAGGTGCAGAAATTCGAGGTCAATTTGGGGGCTACCCGTTCTATC	4837
Db	130	GluGlnArgArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyIysLysSerLeu	149	Db	281	-----	281
QY	3758	GAGCAGTGGGTGACCGAGGAGCGGCTGCTTTGTGCGCTTTCGCCGACCAAGCCGGTG	3817	QY	4838	CCCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGGAGGCCCTGAGAGCAGCTTCAATGA	4897
Db	150	GluGlnTrpValThrGluGluAlaAlaCysLeuCysAlaAlaPheAlaAsn-----	166	Db	282	-----AlaLysGlyAsnProGluSerSerPheAsnAs	292
QY	3818	GGTGATGGGAGAGGACAAAGCGGGAATCGGGAAGCGGGGAGCGGAGAGGCAACC	3877	QY	4898	TGAGACCTGCGCATAGTGGTAACTGTTCTTGGGGATGGTGCACACCTCGAC	4957
Db	166	-----	166	Db	292	pGluAsnLeuArgIleValAlaAlaAspLeuPheSerAlaGlyMetValThrThrSerTh	312
QY	3878	CCTTACCCGATCTCCCGACCCCGAGAGCCCTTTTGGCCCAAGCGCTCTTGACAA	3937	QY	4958	CACGCTGGGCTGGGGCTCTCTCATGATCTTACACCTTGGATGTGTCAGCGTGGGCCAG	5017
Db	167	-----His-Ser-----GlyArgProPheArgProAsnGlyLeuLeuAspLy	180	Db	312	rThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGln-----	328
QY	3938	AGCCGTGAGCAACGTGATCCCTCCTCAGCTCGGGCGCGCTTCGAGTACGACGCC	3997	QY	5018	CTGGGGCCCAAGGAGGAGTGTAGGGAGGAGGTACAGCTGGGGGCGCTGGGCTTACG	5077
Db	180	sAlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTrpAspAspPr	200	Db	328	-----	328
QY	3998	TCCCTTCTCAGCTGTGACCTAGCTCAGGAGGACTGCAAGAGAGTTCGGGCTTCT	4057	QY	5078	TGGGACACCGGGGCTTCCAGCACAGCGCTGGCCAGGCTCTGTAAAGCTTACTTCTCCTC	5137
Db	200	oArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLe	220	Db	328	-----	328
QY	4058	GCSCGAGTCCGAGCGAGAGACCGAGGAGTCTCTGCAGGCGGAGCTCCTGAGAGTGCC	4117	QY	5138	AAACACAGGAGGAGAGTGTCCCTGGGTGTGACCCATTTGGGGGAGCATGTCTGT	5197
Db	220	uArgGlu-----	222	Db	328	-----	328
QY	4118	GGGGCTGGAAGTGGGGCTCCGAAAGGCGAGATTTGCATAGATGGGTTTGGGAAAGCAT	4177	QY	5198	CCAGTCCGTGTCCAAAGAGAGATCGACGACGTGATAGGCGAGGTGGCGCACAGAGATG	5257
Db	222	-----	222	Db	329	-Arg-ArgValGlnGlnGluIleAspAspValIleGlyGlnValArgArgProGluMet	347
QY	4178	TCCAGGAGACCCACTGTAAAGAGGGCTGGAGGAGGAGGGGACATCTCAGACATGGTGC	4237	QY	5258	GGTGACCGGCTCAGATGCCCTTACACCTGCGGTGATTTCAGAGGTGAGCGCTTTGGG	5317
Db	222	-----	222	Db	348	GlyAspGlnAlaHisMetProTyrThrThrAlaValIleHisGluValGlnArgPheGly	367
QY	4238	TGGGAGAGGTGTGCCGGGTGAGGGGCACCGAGAGGCGCAAGGACTCTGTACCCCGT	4297	QY	5318	GACATCATCCCTGAGTGTGACCCATATGACATCCGTGACATCGAAGTACAGGCTTC	5377
Db	222	-----	222	Db	368	AspIleValProLeuGlyMetThrHisMetThrSerArgAspIleGluValGlnGlyPhe	387
QY	4298	CCAGTTTGAGATTTTCGATTTTAGTTTCTCTCTGGGCAAGGAGAGGGTGGAGGCTG	4357	QY	5378	CGCATCCCTTAAGGTAGGCGCTGGCGCCCTCTCACCCAGCTCAGACACCAGACCTGGTGA	5437
Db	222	-----	222	Db	388	ArgIleProLys-----	391
QY	4358	GCACCTGGGAGGGACTTGTGTAGGTCACTGGGTAAAGACAGGCGGCCCTGGGTCTACCT	4417	QY	5438	TAGCCCGACGATGGCTACTGCGAGGTGGGCGCCCACTCTAGGAACCTTGGCCACCTAGTCTC	5497
Db	222	-----	222	Db	391	-----	391
QY	4418	GGAGATGGCTGGGGCTTGAGACTTGTCCAGGTGAACGACAGGACACAGGAGGATTGAGAC	4477	QY	5498	CAATGCCACCACTGACTGTCCCGACTTGGGTGGGGGTCCAGAGATATAGGCAGGCGCTG	5557
Db	222	-----	222	Db	391	-----	391
QY	5558	GCCTGTCCATCCAGAGCCCGCTGTAGTGGGGAGACAAACAGGACCTGCCAGATGTTG	5617	QY			

```
Db 391 ----- 391
QY 5618 GAGGAGCCAGCGCTGACGGAGAGGGGGAGGTGTGGGTGCTCTGAGAGGTGTGACTGC 5677
Db 391 ----- 391
QY 5678 GCCCTGCTGTGGGTGGAGAGGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACA 5737
Db 391 ----- 391
QY 5738 GAGTCAGCTGTGTGCCAGGAGGTGTGTCTCCCGTGTGTGTGGTGGCAGGGTCCACAG 5797
Db 391 ----- 391
QY 5798 CATCTAGAGTCCAGTCCCACTCTCACCTGCATCTCTCGTCCCAAGGAAACACACTCAT 5857
Db 392 ----- 396
QY 5858 CACCAACTGTCTATCGTCTGAGGATGAGCGCTGTGGGAGAGCCCTTCGGTTCCA 5917
Db 396 eThrAsnLeuSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHi 416
QY 5918 CCCCAGAACTTCTCTGATGCCAGGGCCACTTTGTGAAGCCGAGGCTTCTCGCTTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPh 436
QY 5978 CTCAGAGTGTCTGTGGGAGCCCGGCTCTCTGCTCCCTCCCTGGAGTCTTTCAGGGG 6037
Db 436 e--Ser----- 437
QY 6038 TATCACCAGGAGCAGGCTCACTGACGCCCTCCCTCCCAAGGCGCGTGCATGC 6097
Db 438 -----AlaGlyArgArgAlaCys 443
QY 6098 CTCGGGAGCCCTGCGCCGATGAGCTTCTCTTCTTCCCTCCTGCTGTCAGCAC 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThSerLeuLeuGlnHis 463
QY 6158 TTCAGTCTTCTGTCGCGCCCGAGAGCCCGCCGAGCCTCTCTGTCGTCTGCTGCTTT 6217
Db 464 PheSerPheSerValProThrGlyGlnProArgProSerHisGlyValPheAlaPhe 483
QY 6218 CTGCTGACCCATCCCTACAGCT--TTGTCTGTGTGCCCCCTAGATGGGTACC 6273
Db 484 LeuValThrProSerProTyrGluTyrLeuCysCysAlaProLeuGluTrpGlyThr 502

RESULT 2
ABU09594
ID ABU09594 standard; protein; 502 AA.
XX
AC ABU09594;
XX
DT 16-JUL-2003 (first entry)
XX
DE Human cytochrome p450 gene CYP2D6, variant C5816AT, protein.
XX
KW Human; cytochrome P450; CYP2D6; chromosome 22; SNP;
KW single nucleotide polymorphism; drug metabolism; cardiovascular disorder;
KW psychiatric disorder; drug sensitivity.
XX
OS Homo sapiens.
XX
FH Key
FH Location/Qualifiers
FT Misc-difference 492..502
FT /note= "These amino acids replace the wild-type sequence
FT (Leu-Cys-Ala-Val-Pro-Arg) as the result of a single-
FT nucleotide polymorphism which causes a frameshift"
XX
XX
XX EP1281755-A2.
XX
XX 05-FEB-2003.
XX
```

```
PF 16-JUL-2002; 2002EP-00254972.
XX
PR 31-JUL-2001; 2001US-0309111P.
XX
PA (PFIZ ) PFIZER PROD INC.
XX
PI Milos PM, Webb SM;
XX
DR WPI; 2003-373769/36.
DR N-PSDB; ACA61304.
XX
PT New cytochrome P450 2D6 gene variants and polypeptides, useful for
PT determining if a subject has or is at risk of developing a drug
PT sensitivity condition or disorder that is associated with an aberrant
PT CYP2D6 activity.
XX
PS Claim 4; Fig 5; 88pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising a cytochrome
CC P450 2D6 gene variant, e.g. G5799C or C5816AT (referring to the genomic
CC sequence or the same variant nucleotide in the corresponding cDNA
CC sequences). Also included are probes, primers (allele specific
CC oligonucleotides) and arrays used to detect and/or amplify the CYP2D6
CC gene polymorphic regions, the variant polypeptides, antibodies which are
CC capable of distinguishing between the variant and wild-type polypeptides,
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
CC determining whether an individual is susceptible to being a poor
CC metaboliser of drugs. The DNA probe is useful for hybridising to a
CC variant form of the CYP2D6 gene. The primer is useful for amplifying the
CC C5816TA allelic variant. The allele specific nucleotide is useful for the
CC detection of the C5816TA allelic variant. The methods are useful for
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and
CC determining if an individual is susceptible to being a poor metaboliser
CC of drugs. The nucleic acids are useful as probes or primers for
CC determining whether a subject has a genetic deficiency for metabolising
CC drugs that are substrates of P450 CYP2D6. The methods are useful for
CC determining if a subject has or is at risk of developing a drug
CC sensitivity condition or disorder that is associated with an aberrant
CC CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
CC aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
CC appropriate drugs or determining the course of treatment to administer to
CC a subject to treat cardiovascular or psychiatric disorders, or for
CC treating a subject with a drug sensitivity or disorder associated with a
CC specific allelic variant of a polymorphic region of the CYP2D6 gene. The
CC antibodies are useful for monitoring CYP2D6 protein levels in an
CC individual for determining whether a subject has a disease or conditions
CC associated with an aberrant CYP2D6 protein level. The gene is located on
CC human chromosome 22. The present sequence is the C5816AT variant CYP2D6
CC protein
XX
SQ Sequence 502 AA;
XX
Alignment Scores:
Pred. No.: 9,54e-107 Length: 502
Score: 1803.00 Matches: 468
Percent Similarity: 34.31% Conservative: 14
Best Local Similarity: 33.31% Mismatches: 20
Query Match: 9.63% Indels: 904
DB: Gaps: 9
US-09-820-788A-3 (1-10278) x ABU09594 (1-502)
QY 2078 ATGGGGCTAGAACGACCTGGTGGCCCTGGCCATGATAGTGGCCATCTTCTCTCTGCTG 2137
Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
QY 2138 GACCTGATGACCGGCACCAACGCTGGCTGCAGCTACCCGCCAGGTCCCTGCCACTG 2197
Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaAlaArgTyrProGlyProLeuProLeu 40
QY 2198 CCCGGCTGGCAACCTTGTGCTGTGGACTTCCAGAACACACCATACTGCTTCGACCA 2257
```


Db 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyr-CysPheAspG1 60
2258 GGTGAGGAGGAGGTCTCTGAGGGCGCAGAGGTCTCTGAGGATGCCCCACCACAGCAAA 2317
Db 60 n----- 60
2318 CATGGTGGTGGTTAAACACAGCTGGATCAGAACCCAGGCTGAGAGGGGAGCAGG 2377
Db 60 ----- 60
2378 TTTGGGGACCTTCCTGGGGAAGACATTTATACATGGCATGAAGACTGGATTTTCCAA 2437
Db 60 ----- 60
2438 AGCCCAAGAGAGTAGGGCAAGGCTGGAGGTGGAGCTGGACTTTGGCAGTGGGCATGC 2497
Db 60 ----- 60
2498 AAGCCCATTTGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGTGTGACACCAAGGA 2557
Db 60 ----- 60
2558 AAGGCCTTGGGAATGGAAGTAGTGTAGTCTGAGTGCCTGTTAAATCAGAAATCGAGG 2617
Db 60 ----- 60
2618 ATGAAGGGGTGCAGTGACCCGGTTCAAACTTTTGCACCTGTGGTCTCTGGGCCTCACT 2677
Db 60 ----- 60
2678 GCTCACCGGCATGGACCATCATCTGGGAATGGATGCTAACTGGGGCCTCTCGGCAATTT 2737
Db 60 ----- 60
2738 TGGTGACTCTTCAAGGTCAATACCTGGGTGACGATCCAAACTGAGTTCTCCATCACAG 2797
Db 60 ----- 60
2798 AAGGTGTGACCCCAACCTTGCCCCAGATCAGAGGTGGGTCTCTCTTCCACCTGC 2857
Db 60 ----- 60
2858 TCACTCTGTAGTACCCCGGGGTCTGTCGAAGTTCAATAGGACTAGGACCTGTAGTCTG 2917
Db 60 ----- 60
2918 GGTGTATCTGTGCTTGACAAGAGCCCTGACCTCTCTGAGTTTGGCGCGCTTCTG 2977
Db 61 -----Leu-Arg-ArgPheG 66
2978 GGGACGTGTTACCTGACGTGGCTTGGACGCGGTGTGTGCTCAATGGGCTGGCGG 3037
Db 66 llyAspValPheSerLeuGlnLeuAlaTrpThrProValValLeuAsnGlyLeuAla 86
3038 CCGTGGCGAGGCGATGTGTACCCGCGGAGACACGCGCCAGCCGCGCTGGCCCA 3097
Db 86 llaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProValProI 106
3098 TCTACAGGTCTCTGGGCTTGGGCGCGCTTCCCAAGGCAAGCGGTGGGGGACAGAGA 3157
Db 106 lethrGlnleLeuGlyPheGlyProArgSer-Gln----- 117
3158 CCGCGTTCTCGTGGGCGCGGTGGACAGTGACCTGACCTAGCCCAAGACGCGCAGAGGGG 3217
Db 117 ----- 117
3218 TGGGTCCTGGACGTGAACAGAGATAAAGGCCAGCGAGTGGCTGAGGACAGTGGGCCA 3277
Db 117 ----- 117
3278 GGAACCACTGCACTGGGGAGGTGCGAGTCTGTGGGCTGGGAGGGGCGGCTACTGC 3337

117 ----- 117
3338 CCAGACCCGCCAGAAAGCCCGGTGGCGAGGCTGATGCGTGAAGTGGCGGTGGCGGGAC 3397
117 ----- 117
3398 CCGCGCTATGCTGCGGGCTCACTGTGGCGGACCGGGCGGATCTTCTTTGAGTGAAAG 3457
117 ----- 117
3458 GTGGTCAAGGTGGCGAGACGAGGTGGGGCCAAACCCCGCCAGGAGGGAGCAATG 3517
117 ----- 117
3518 TGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGCTAGGACTGCGGGAGACC 3577
117 ----- 117
3578 TTTGGAGCGCCAGGCTTGGAGTGGGTGGCGAGGTGGGCCAAGGCTTCTATGGCAAC 3637
117 ----- 117
3638 GCCACAGTGTCCGTCCCGCCCCAGGGGTGATCTGTGCGCTATGGGCCCGGTGGCGC 3697
118 -----GlyValPheLeuAlaArgTyrGlyProAlaTrpArg 129
3698 GAGCAGAGCGCTTCTCCGTGTCCACCTTGGCGAACTTGGGCTGGGCAAGTGCCTG 3757
130 GluGlnArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysSerLeu 149
3758 GAGCAGTGGGTGACCGAGGAGCGCTTGTGCGCTTGTGCGGCTTCCCGCAACGCGGTG 3817
150 GluGlnTrpValThrGluGluAlaAlaCysLeuCysAlaAlaPheAlaAsn----- 166
3818 GGTGATGGCAGAGGGCACAAAGCGGAACTGGGAAGCGGGGAGCGAGAGGCAACC 3877
166 ----- 166
3878 CCTTACCCGCATCTCCCAACCCCGAGGAGCGCCCTTTCGCCCAACGGCTCTTGACAA 3937
167 -----His-Ser-----GlyArgProPheArgProAsnGlyLeuLeuAsp 180
3938 AGCCGTGAGCAACGTGATCGCTCTCCTCACCCTGGCGGCGCTTTCGAGTACGACACC 3997
180 sllaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAsp 200
3998 TCGCTTCTCAGGCTGTGGACTAGCTCAGGAGGACTGAAGAGGAGTTCGGGCTTCT 4057
200 oArgPheLeuArgLeuLeuAspLeuAlaGlnGlyLeuLysGluGluSerGlyPheLe 220
4058 GCGCAGGTGGCGGAGGAGACCGAGGAGTCTCTGCGGGGCGAGCTCTGAGAGGTGCC 4117
220 uArgGlu----- 222
4118 GGGGTGAGCTGGGCGCTCCGAAGGCGAGGATTTGCATAGATGGGTTTGGGAAAGCAT 4177
222 ----- 222
4178 TCCAGGAGACCCCACTGTAAGAAGGCGCTGGAGGAGGAGGAGCATCTCAGACATGCTG 4237
222 ----- 222
4238 TGGGAGAGTGTGCCCGGTGAGGGGCGACAGGAGAGGCCAAGGACTCTGTACCCCTGT 4297
222 ----- 222
4298 CCACGTGGAGATTTTCGATTTTAGGTTTCTCTCTGGGCAAGGAGAGGAGGCTG 4357
222 ----- 222
4358 GCACCTGGGAGGAGCTTGGTGGTGTAGTGTAGTGTAGGAGGAGGAGGAGGCTGTACCT 4417
222 ----- 222

```
QY 4418 GGAGATGGCTGGGGCTGAGACTTGTCCAGGTGAACGACAGACACAGGAGGATTTGAGAC 4477
Db 222 ----- 222
QY 4478 CCCGTTCTGTCTGGTGTAGTCTGAATGCTCTCCCGTCTCTCCGACATCCACGCGT 4537
Db 223 -----ValLeuAsnAlaValProValLeuLeuHisIleProAlaIle 236
QY 4538 GCGTGGCAAGTCTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGAGCTCTTAAC 4597
Db 236 uAlaGlyValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TCAGCACAGGATGACTGGGACCCAGCCAGCCAGCCAGCCAGCTGAGCGCTTCTCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLeu 276
QY 4658 GCCTAAGAGGAGAGGTGAGAGTGGCTGCCACGCTGGGGGCAAGGGTGGTGGTTGAA 4717
Db 276 uAlaGluMetGluLys----- 281
QY 4718 CGTCCCAGGAGGAATGAGGGGAGGCTGGGCAAAAGGTTGGACCACTGATCACCCGGCGA 4777
Db 281 ----- 281
QY 4778 GCCGCATCTGGGCTGACAGGTGCAGAAATTGGAGGTCAATTGGGGGCTACCCCGTTCTATC 4837
Db 281 ----- 281
QY 4838 CCTGAGTATCTCTCGGCCCTGTCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCAATGA 4897
Db 282 -----AlaLysGlyAsnProGluSerSerPheAsnAs 292
QY 4898 TGAGAACCTGCCATAGTGTGGGTAACTGTCTTCTGCGGGATGGTGACCACTCCGAC 4957
Db 292 pGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetValThrThrSerTh 312
QY 4958 CACGCTGGCTGGGCTCTCTGCTCATGATCTCATCTGATGATGCGAGCGTGACCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGln----- 328
QY 5018 CTGGGGCCCAAGCAGGACTGAGGGAGGAGGTACAGCTGGGGGCCCCCTGGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCCGGGGCTTCAGACACAGGCGTGCCAGGCTCCTGTAAAGCCTAACTTCCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGGAGGAAGGAGAGTGTCCCTGGGTGCTGACCCATTTGTGGGACGCGATCTCTGT 5197
Db 328 ----- 328
QY 5198 CCAGTCCGCTGTCCACAGGAGATGACGACCGTGTAGGCGAGGTGCGGCGACAGAGATG 5257
Db 329 -Arg--ArgValGlnGlnIleAspValIleGlyGlnValArgArgProGluMet 347
QY 5258 GGTGACCAAGGCTCACATGCCCTTACACCACTGCGGTGATTCACAGGTGACGCGTTTGGG 5317
Db 348 GlyAspGlnAlaHisMetProTyrThrThrAlaValIleHisGluValGlnArgPheGly 367
QY 5318 GACATCATCCCTGAGTGTGACCCATGACATCCGTCGATCGAAGTACAGGCTTC 5377
Db 368 AspIleValProLeuGlyMetThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CGCATCCCTTAAGTAGGCTGGCGGCCCTCTCACCCAGCTCAGCACAGCAGCTGGTGA 5437
Db 388 ArgIleProLys----- 391
QY 5438 TAGCCCCAGCATGGGTACTGCGAGGTGGGCCCACTCTAGGAACCTGGCCACCTAGTCTCT 5497
Db 391 ----- 391
```

```
QY 5498 CAATGCCACCACACTGACTGTCTCCCACTTGGTGGGGGGTCCAGAGTATAGCAGGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTCCATCCAGAGCCCCCGTCTAGTGGGGGAGACAAACAGGACCTGCCAGATGTTG 5617
Db 391 ----- 391
QY 5618 GAGNACCCAGCGCTGACAGGAGAGGGGGCAGTGTGGTGCCTCTGAGAGGTGTGACTGC 5677
Db 391 ----- 391
QY 5678 GCCCTGTGTGGGGTTCGAGAGAGGTACTGTGGAGCTTCTCGGGGCGCAGGACTAGTTGACA 5737
Db 391 ----- 391
QY 5738 GAGTCCAGCTGTGTGCCAGGCAGTGTGTGTCTCCCGTGTGTGTGGTGGCAGGGGTCCAG 5797
Db 391 ----- 391
QY 5798 CATCTAGATGTCAGTCCCACTCTCACCTGTCATCTCTGCCAGGGAACGACACTCAT 5857
Db 392 -----Gly-ThrThrLeuIle 396
QY 5858 CACCAACCTGTCATCGGTGCTGAAGGATGAGCGCTCTGGGAGAGCCCTTCCCGTTCCA 5917
Db 396 eThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHi 416
QY 5918 CCCCAGAACACTTCTTGATGCCAGGGCCACTTTTGAAGCCGGAGGCTTCTGCTCTTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPh 436
QY 5978 CTCAGCAGTGCCTGTGGGGAGCCGGCTCCTCTGCTCCCTCCGTGGAGTCTTGACAGGG 6037
Db 436 e--Ser----- 437
QY 6038 TATCACCCAGGAGCCAGGCTCACTGAGCGCCCTCCCTCCACAGGCGCGCTGCATGC 6097
Db 438 -----AlaGlyArgArgAlaCys 443
QY 6098 CTCGGGAGCCCCCTGGCCCGCATGAGCTTCTCTTTCACCTCCCTCTCTGCAGCAC 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeuLeuGlnHis 463
QY 6158 TTCAGCTTCTCCGTGGCGCCGCGAGCCCGGCGAGCCCGGCGCCAGCACTCTCTGTCTGCTAGCTTT 6217
Db 464 PheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyValPheAlaPhe 483
QY 6218 CTGTGTGACCCCATCCCTACGAGCT-TTGTGTGTGTGCCCGCTAGATGGGTACC 6273
Db 484 LeuValSerProSerProTyrGluTyrLeuCysAlaProLeuGluTrpGlyThr 502

RESULT 3
AAR72376
ID AAR72376 standard; protein; 497 AA.
XX
AC AAR72376;
XX
DT 25-MAR-2003 (revised)
DT 15-NOV-1995 (first entry)
XX
DE Human auxiliary cytochrome P450 species 2D6 variant 1 protein.
XX
KW Human cytochrome P450; amplification; PCR; primer; expression vector;
KW yeast NADPH-P450 reductase; safety; fusion protein; metabolite;
KW carcinogen; mutagen; liver metabolism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 296
FT /note= "Cys to Arg variation"
XX
```

PN EP644267-A2.
XX 22-MAR-1995.
XX 20-JUL-1994; 94EP-00111298.
XX 20-JUL-1993; 93JP-00201120.
PR 21-JUL-1993; 93JP-00180246.
PR 30-JUL-1993; 93JP-00208279.
XX (SUMO) SUMITOMO CHEM CO LTD.
XX Hayaashi K, Sakaki T, Yabusaki Y, Komai K, Kaneko H, Nakatsuka I;
PI WPI; 1995-116991/16.
XX N-PSDB; AAQ87730.
DR Evaluation of safety of a chemical cpd. - using recombinant yeast
XX expressing human cytochrome p450 and a yeast NADPH-P450 reductase.
PT Example; Page 87-89; 124pp; English.
XX The amino acid sequence of the human auxillary cytochrome P450 species
CC 2D6 variant 1. This variant contains a variation at residue 296: Cys to
CC Arg, caused by a variation at base 886: T to C in the DNA sequence. The
CC cDNA was amplified by PCR using the primers AAQ87763-6. The product was
CC cloned into the yeast expression vectors pAAH5N or pAHR to produce the
CC vectors p2D6 variant 1 for the expression of the cytochrome P450 alone or
CC p2D6R variant 1 for co-expression with the yeast NADPH-P450 reductase.
CC The vectors are used in a method for evaluating the safety of a chemical
CC compound by reacting the chemical compound with recombinantly produced
CC human cytochrome P450 molecular species 1A2 (AAQ87714), 2C9 (AAQ87715),
CC 2E1 (AAQ87716), or 3A4 (AAQ87718), or their auxillary species and
CC variants (AAQ87718-32), and yeast NADPH-P450 reductase, either as a fused
CC protein or in cell extracts, and analysing the resulting metabolite to
CC assess the safety of the chemical compound. The method is useful for
CC determining whether the chemical compound, or its metabolite, will be
CC converted into a carcinogenic or mutagenic form through metabolism in the
CC liver. (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 497 AA;
SQ
Alignment Scores:
Pred. No.: 1.1e-106 Length: 497
Score: 1802.00 Matches: 466
Percent Similarity: 34.14% Conservative: 12
Best Local Similarity: 33.29% Mismatches: 19
Query Match: 9.62% Indels: 903
DB: 2 Gaps: 9
US-09-820-788A-3 (1-10278) x AAR72376 (1-497)
QY 2078 ATGGGGCTAGAGCACTGGTGGCCCTGGCAGTATAGTGGCCATCTTCCTGCTCTGGTG 2137
DB 1 MetGlyLeuGluAlaLeuValProLeuAlaValAlaLeuValAlaLeuLeuVal 20
QY 2138 GACCTGATGACCGGCACCAACGCTGGGTGACGCTACCCGCGAGTCCCTGCCACTG 2197
DB 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgGlyProProGlyProLeuVal 40
QY 2198 CCCGGCTGGGCAACCTTCTGCTGATGGGACTTCCAGAACACACCACTACTGCTTGACCA 2257
DB 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspG 60
QY 2258 GGTAGGGAGGAGGTCTCTGGAGGGCGGAGAGTCTCTGAGGATGCCCCACCCAGCAMA 2317
DB 60 n----- 60
QY 2318 CATGGGTGGTAAACACAGGCTGGATCAGAGCCAGGCTGAGAGGGGAGCAGG 2377
DB 60 ----- 60
QY 2378 TTTGGGGAGCTTCTCTGGGGAAGGACATTTATACATGCGATGAAGGACTGGATTTTCCAA 2437

Db 60 ----- 60
QY 2438 AGGCCAAGGAAGAGTAGGCAAGGGCTTGGAGTGGAGCTTGGAGTGGAGTGGCATGC 2497
Db 60 ----- 60
QY 2498 AAGCCCATTTGGGCAACATATATTATATGAGTACAAAGTCCCTTCTGTGTGACACCAAGGA 2557
Db 60 ----- 60
QY 2558 AAGGCTTTGGGAATGGAAGATGAGTTAGTCTCTGAGTGGCGCTTTAAATCATCAAGAAATCGAGG 2617
Db 60 ----- 60
QY 2618 ATGAAGGGGTGACGTGACCCGGTTCAAACTTTTGCACGTGTGGGTCTCTCGGCCCTCACT 2677
Db 60 ----- 60
QY 2678 GCTCACCGCATGGACCATCATCTGGGAATGGATGCTAACTGGGGCTCTCGGCAATTT 2737
Db 60 ----- 60
QY 2738 TGGTGACTCTTCAAGGTCTATACCTGGGTGACGCATCCAAACTGAGTTCTCTCATCACAG 2797
Db 60 ----- 60
QY 2798 AAGGTGTGACCCCTGCTGCCCATCAGATCAGGAGCTGGGTCTCTCTCTCCACCTGC 2857
Db 60 ----- 60
QY 2858 TCACCTCTGTAGTCCCGGGGTCTGTCAAAGTTCAAATAGGACTAGGACCTGTAGTCTG 2917
Db 60 ----- 60
QY 2918 GGGTGATCTGTGCTTGACAAGAGGCCCTGACCCCTCTCTGAGTTGCGGCGCGCTTTCG 2977
Db 61 ----- 66
QY 2978 GGGAGCTTTTCAAGCTTGGCTGGAGTGGCTGGAGTGGCTGGCTGGCTGGCTGGCTGGCTGG 3037
Db 66 lYAspValPheSerLeuGlnLeuAlaTrpThrProValValLeuAenGlyLeuAla 86
QY 3038 CGTGGCGAGCGGATGGTGGACCCCGCGGAGGACACACCGCCGCGCTGGCTGGCTGGCTGG 3097
Db 86 lAlaValArgGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProProValPro 106
QY 3098 TCTACAGGTCTCTGGCTTGGGCGCGCTTCCCAAGGCAAGCGCGGTGGGGGACAGAGA 3157
Db 106 lThrGlnIleLeuGlyPheGlyProArgSer-Gln----- 117
QY 3158 CGCGCTTTCCGTGGGCGCGGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 3217
Db 117 ----- 117
QY 3218 TGGGGTCTTGGACGTGAAACAGAGATAAAGGCGAGCGAGTGGGCTGAGGACAGTGGGCGCA 3277
Db 117 ----- 117
QY 3278 GGAAACCACTGTGACCGGGGAGGTGCGAGTCTGTGGGTGGGAGGGGGCGGGGCTACTGC 3337
Db 117 ----- 117
QY 3338 CCAGACCCGCCAGAACCCGGTGGGGAGGCTGATCGCTGAAGTGGCGGTGGCGGGGAC 3397
Db 117 ----- 117
QY 3398 CGCGCTATGTGCGGGCTCAGTGTGGGGGAGCGCGGGGATCTTCTTGTAGTGAAG 3457
Db 117 ----- 117
QY 3458 GTGGTGGGGGAGAGACGAGGTGGGGCCAAACCCCGCCCGAGGAGGGGAGCAATG 3517

Db 117 ----- 117
Qy 3518 TGGGTGAGCAAGAGTGGSCCTGTGCCCAGCTGACCGGGCTAGGGACTGGCGGAGACC 3577
Db 117 ----- 117
Qy 3578 TTGTGAGCGCCAGGCTTGGAGTGGGTGGCGGAGGTGGGCCAAGGCTTTCATGGCAAC 3637
Db 117 ----- 117
Qy 3638 GCCACAGTGTCCGTCCGCCGCCAGGGGTGATCCTGTGCGGTATGGGCGCGCGTGGCGC 3697
Db 118 ----- -GlyValPheLeuAlaArgTyrGlyProAlaTrpArg 129
Qy 3698 GAGCAGAGCGCTTCTCGTGTCCACCTTGGCACTTGGGCTGGGCAAGAGTCGCTG 3757
Db 130 GluGlnA-ArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysSerLeu 149
Qy 3758 GAGCAGTGGGTGACCGAGAGCGCCCTGCTTGTGCGCTTTCGCCGACCAAGCCGGTG 3817
Db 150 GluGlnTrpValThrGluAlaAlaCysLeuCysAlaAlaPheAlaAsn----- 166
Qy 3818 GGTGATGGGCAGAGGGCACAAGCGGAACTGGGAAGCGGGGACGGAGAGGCAACC 3877
Db 166 ----- 166
Qy 3878 CCTTACCGCATCTCCCAACCCCGAGAGCGCCCTTCGCCCAACGGCTCTTCGACAA 3937
Db 167 -----His-Ser-----GlyArgProPheArgProAsnGlyLeuLeuAspLys 180
Qy 3938 AGCCGTGAGCAACGTGATCGCTCCCTCACCTGCGGGCGCCCTTCGAGTACGACGCC 3997
Db 180 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPr 200
Qy 3998 TGGCTTCTCAGGCTGTGGACCTAGCTCAGAGGAGCTGAAGAGGAGTGGGCTTTCT 4057
Db 200 oArgPheLeuArgLeuLeuAspLeuAlaGlnGlyLeuLysGluSerGlyPheLe 220
Qy 4058 GCGCAGGTGCGGAGCGAGAGACCAGGAGTCTCTGCGGGGCGAGCTCTGAGAGGTGCC 4117
Db 220 uArgGlu----- 222
Qy 4118 GGGGTGAGCTGGGGCTCCGAAGGCGAGGATTTGCATAGATGGGTTTGGGAAAGACAT 4177
Db 222 ----- 222
Qy 4178 TCCAGGAGACCCCACTGTAAAGGGCCTGGAGAGGAGGGGACATCTCAGACATGGTGG 4237
Db 222 ----- 222
Qy 4238 TGGGAGAGGTGTGCCGGGTGAGGGGCGACAGAGAGGCCAAGACTCTGTACCCCGT 4297
Db 222 ----- 222
Qy 4298 CCACGTGGAGATTTTCGATTTTGGTTTCTCTCTGCGCAAGGAGAGGGGTGGAGGCTG 4357
Db 222 ----- 222
Qy 4358 GCATTTGGGGAGGGACTTGGTGAGGTGAGTGAAGGACAGGCGCCCTGGGTCTACCT 4417
Db 222 ----- 222
Qy 4418 GGAGATGGCTGGGGCTGAGACTTGTCCAGGTGAACGCGAGAGCAGGAGGATTTGAGAC 4477
Db 222 ----- 222
Qy 4478 CCCGTTCTGTGTGTAGTGTGATGCTGATGCTGCTCCCGTCTCTGTCACATCCAGCGCT 4537
Db 223 -----ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe 236
Qy 4538 GCGTGGCAAGGTCTACGCTTCCAAAGCTTTCTGACCCAGCTGGATGAGCTGCTAAC 4597
Db 236 uAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuLeuAspGluLeuTh 256

Qy 4598 TGAGCACAGGATGACTTGGGACCCAGCCAGCCAGCCAGACCTGACTGAGGCTTCTCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
Qy 4658 GGCAAAGAGAGAGAGTGTGAGTGGCTGCGCACCGTGGGGGCAAGGGTGTGGTTGAA 4717
Db 276 uAlaGluMetGluLys----- 281
Qy 4718 CGTCCCAGGGAATGAGGGGAGGCTGGGCAAAAGTTGGACCAGTGCATCACCCGCGCA 4777
Db 281 ----- 281
Qy 4778 GCCGATCTGGCTGCACAGGTGCAGAAATTGGAGGTCAATTGGGGGTACCCGTTCTATC 4837
Db 281 ----- 281
Qy 4838 CCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGACAGCTTCAATGA 4897
Db 282 -----AlaLysGlyAsnProGluSerSerPheAsnAs 292
Qy 4898 TGAGAACCTGGCATAGTGGTGGTAACTGTTCTTCGCGGATGGTGCACACCTCGAC 4957
Db 292 pGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetValThrThrSerTh 312
Qy 4958 CACGTGCGCTGGGGCTCTCTGCTCATGATCTACACCTGATGTGCAGCGTGCAGCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuMetIleLeuHisProAspValGln----- 328
Qy 5018 CTGGGGCCCAAGGCAGGAGCTGAGGGAGGAAGGTACAGTGGGGGCCCCCTGGGCTTACG 5077
Db 328 ----- 328
Qy 5078 TGGGACACCCGGGGCTTCCAGCA CAGGCGTGGCCAGGCTCTGTGAAGCTTAATCTCTCC 5137
Db 328 ----- 328
Qy 5138 AACACAGGAGGAAGAGAGTGTCCCTGGTGTGACCCATTGTGGGAGCGCATGCTGT 5197
Db 328 ----- 328
Qy 5198 CCAGTCTGTGTCACAGGAGATCGACAGCTGATAGGCGAGGTGCGGCGACAGAGATG 5257
Db 329 -Arg-ArgValGlnGlnIleAspValIleGlyGlnValArgArgProGluMet 347
Qy 5258 GGTGACAGGCTCACATGCTCTACACCTGCGTGTATTACAGGTGCAGCGCTTTGGG 5317
Db 348 GlyAspGlnAlaHisMetProTyrThrThrAlaValIleHisGluValGlnArgPheGly 367
Qy 5318 GACATATCCCTCAGTGTGACCATATGACATCCCTGATCATCGAAGTACAGGCTTC 5377
Db 368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
Qy 5378 GGCATCCCTAAGTAGGCTTGGCGCTCTCACCAGCTCAGCCAGCAGCAGCTGGTGA 5437
Db 388 ArgIleProLys----- 391
Qy 5438 TAGCCCCAGCATGGCTACTGCGAGGTGGGCCCACTCTAGGAACCCCTGGCCACCTAGTCT 5497
Db 391 ----- 391
Qy 5498 CAATGCCACCACACTGCTGTCTCCCACTTGGTGGGGGTCCAGAGTATAGCAGGGCTG 5557
Db 391 ----- 391
Qy 5558 GCCTGTCCATCCAGAGCCCCCTCTAGTGGGGAGACAAACAGGACCTGCCAGAAATGTTG 5617
Db 391 ----- 391
Qy 5618 GAGNCCAGCGCTGAGGAGAGGGGGCAGTGTGGGTGCTCTGAGAGGTGTGACTGC 5677
Db 391 ----- 391

QY 5678 GCCCTGCTGGGGTGGAGAGGGTACTGTGGAGCTTCTCGGCGCAGAGACTAGTTGACA 5737
Db 391 ----- 391
QY 5738 GAGTCCAGCTGTGTGCCAGGACGTGTGTGTCCTCCCGTGTGTTGGTGGCAGGGTCCCGAG 5797
Db 391 ----- 391
QY 5798 CATCTTAGAGTCCAGTCCCACTCTCACCTGCATCTCTGCCAGGAGACGACACTCAT 5857
Db 392 -----Gly-ThrThrLeuI 396
QY 5858 CACCAACTGTATCGGTGCTGAAGATGAGGCGCTGTGGAGAACCTTCGCTTCCA 5917
Db 396 eThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHi 416
QY 5918 CCCCGAACACTTCTCGATGCCAGGCGCACTTTGTGAAGCGGAGCGCTTCTCGCTTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProH 436
QY 5978 CTGACGAGGTGCTGTGTGGGAGCCCGCTCTCCCTGTCCTCCCTTCCGTCGAGTCTTGCAGGGG 6037
Db 436 e-Ser----- 437
QY 6038 TATCACCCAGGAGCCAGGCTCACTGAGCGCCCTCCCTCCCAAGCGCGCGTGCATGC 6097
Db 438 -----AlaGlyArgArgAlaCys 443
QY 6098 CTCGGGGAGCCCTGGCCGCGCATGGAGCTTCTCTTACCTCCCTGCTGCAGGAC 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeuLeuGlnHis 463
QY 6158 TTCAGCTTCTCCGTGGCGCGGACAGCCCGCGCCAGCCACTCTCTGTCGTCAGCTTT 6217
Db 464 PheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyValPheAlaPhe 483
QY 6218 CTGGTACCCCACTCCCTCCCTACGAGCTTTGTGCTGTCGCCCGC 6259
Db 484 LeuValThrProSerProTyrGluLeuCysAlaValProArg 497
RESULT 4
AAR93183
ID AAR93183 standard; protein; 497 AA.
AAR93183;
11-OCT-1996 (first entry)
Human cytochrome P450 molecular species 2D6 variant #2 protein.
Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;
liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;
evaluation; safety; fusion protein; metabolite; detoxification;
carcinogenic.
Homo sapiens.
JP08056695-A.
05-MAR-1996.
15-JUL-1994; 94JP-00164184.
20-JUL-1993; 93JP-00201120.
30-JUL-1993; 93JP-00208279.
17-JUN-1994; 94JP-00136053.
(SUMO) SUMITOMO CHEM CO LTD.
WPI; 1996-182311/19.
N-PSDB; AAT28396.
Novel method for the evaluation of the safety of a cpd. - using a human

PT cytochrome P450 and yeast NADPH reductase to determine whether the
PT analyte cpd. is detoxified or metabolised to a carcinogen.
XX
XX
XX Example 1; Page 51-53; 74pp; Japanese.
CC This is the amino acid sequence of the human cytochrome P450 molecular
CC species 2D6 variant #2 protein. The corresp. gene was amplified from a
CC human liver derived cDNA library as 2 fragments of 0.4 and 0.9 kb using
CC primers AAT26953-6. The prod. was cloned into the yeast expression vector
CC pAAH5N to generate plasmid p2D6 for prodn. of the cytochrome only or into
CC the vector pAHR to generate the plasmid p2D6R for co-prodn. with the
CC yeast NADPH-P450 reductase. The sequence is placed under control of the
CC yeast ADH gene promoter and terminator. The vectors are used in a method
CC for evaluating the safety of a cpd. by reacting the test cpd. with
CC recombinantly produced human cytochrome P450 mol. species 1A2 (AAT28380),
CC 2C9 (AAT28381), 2E1 (AAT28382), 3A4 (AAT28383) or their variants
CC (AAT28384-98) together with yeast NADPH-P450 reductase (either as a fused
CC protein or as a cell extract) and analysing the resultant metabolite. The
CC cpd. is considered "safe" if it is detoxified or not rendered
CC carcinogenic or "unsafe" if it is not detoxified or is metabolised to a
CC carcinogenic cpd
XX
XX Sequence 497 AA;
SQ
Alignment Scores:
Pred. No.: 1.1e-106 Length: 497
Score: 1802.00 Matches: 466
Percent Similarity: 34.14% Conservative: 12
Best Local Similarity: 33.23% Mismatches: 19
Query Match: 9.62% Indels: 903
DB: 2 Gaps: 9
US-09-820-788A-3 (1-10278) x AAR93183 (1-497)
QY 2078 ATGGGGCTAGAGAGACTGTGTGCCCTGGCCATGATAGTGGCCATCTTCTGCTCTGGTG 2137
Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
QY 2138 GACCTGATGCACCGCACCAACGCTGGGCTGCACCTACCCGCCAGTCCCTGCCACATG 2197
Db 21 AspLeuMetHisArgArgGlnArgIrpAlaAlaGlyProProGlyProLeuProLeu 40
QY 2198 CCCGGGCTGGGCAACCTTGTCTCATGTGGACTTCCAGAACACACCATCTGCTTCCACCA 2257
Db 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspGI 60
QY 2258 GGTGAGGGAGAGGTCTCTGGAGGGCGGAGAGGTCTCTGAGGATGCCCCACCACCAAAA 2317
Db 60 n----- 60
QY 2318 CATGGTGTGGGTTAAACCACAGGCTGGATCAGAAGCCAGGCTGAGAGGGGAGCAGG 2377
Db 60 ----- 60
QY 2378 TTTGGGGGACGTTCTCTGGGGAGGACATTTATACATGTCATGAAGGACTGGATTTTCCAA 2437
Db 60 ----- 60
QY 2438 AGGCCAAGAAGAGTAGGGCAAGGGCCCTGGAGGTGGAGCTGGAGTGGGGCATGC 2497
Db 60 ----- 60
QY 2498 AAGCCCATTTGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACACGAGGA 2557
Db 60 ----- 60
QY 2558 AAGGCTTTGGAAATGGAAGATGAGTTAGTCTCTGAGTGGCGTTTAAATCAGAAATCGAG 2617
Db 60 ----- 60
QY 2618 ATGAAGGGGTGACGTGACCCCGGTTCAAACTTTTGCACGTGGGTCTCGGGCCTCACT 2677
Db 60 ----- 60

QY 2678 GCTCACCAGCATGAGCATCATCTGGGAATGGATGCTAACTGGGGCCTCTCGGCAATT 2737
Db 60 ----- 60
QY 2738 TGGTGACTTTGCAAGGTATACCTGGGTGACGCATCCAACTGAGTTCTCCATCACAG 2797
Db 60 ----- 60
QY 2798 AAGGTGTACCCCAACCCCTGCCCCACGATCAGGAGGTGGGTCTCTCTTCCACCTGC 2857
Db 60 ----- 60
QY 2858 TCACCTCTGGTAGCCCGGGGTCTGTCAAAGTTCAAATAGACTAGGACCTGTAGTCTG 2917
Db 60 ----- 60
QY 2918 GGGTGATCCTGGCTTTGACAAAGAGCCCTGACCCCTCTGAGTTGGGGCCGCTTCG 2977
Db 61 ----- Leu-ArgArgArgPheG 66
QY 2978 GGGAGCTGTTCAGCTGACGTGGCTGGACGCCGGTGGTCTGCTCAATGGGCTGGCGG 3037
Db 66 lyAspValPheSerLeuGlnLeuAlaIatPThrProValValLeuAsnGlyLeuAlaA 86
QY 3038 CGTGGCGGAGCGATGTTGACCCGCGGAGGACACGCCCGCCGCTGGGCCCA 3097
Db 86 laValargGluAlaLeuValThrHisGlyAspThrAlaAspArgProProValProI 106
QY 3098 TCTACAGTCTCGGCTTCGGCGCGCTTCCAAAGCAAGCGCGGTGGGGGACAGAGA 3157
Db 106 leThrGlnIleLeuGlyPheGlyProArgSer-Gln- 117
QY 3158 CCGCGTTCCGTGGGCGCGGTGGACAGTGCAGTGCAGCCAGCAGCGCCGACAGGGG 3217
Db 117 ----- 117
QY 3218 TGGGCTCTGAGCTGAACAGAGATAAAGCCAGCGAGTGGCTGAGGACAGTGGCCA 3277
Db 117 ----- 117
QY 3278 GGAACACCACTGCAGCGGGAGGTGCGAGTCTGTGGGTGGGAGGGCGGGCTACTGC 3337
Db 117 ----- 117
QY 3338 CCAGACCCCAAGACCCCGGTGGCGAGGCTGATGCTGAAGTGGCGGTGGCGGGGAC 3397
Db 117 ----- 117
QY 3398 CGCGCTATGTCGGGCTCAGTGTGGCGGACGGCGGGATCTTCTTGAGTGAAG 3457
Db 117 ----- 117
QY 3458 GTGGTCAGGGTGGCAGACAGAGTGGGGCCAAACCCCGCCAGGCGGGGAGCAATG 3517
Db 117 ----- 117
QY 3518 TGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGACCGGGCTAGGACTGCGGGAGACC 3577
Db 117 ----- 117
QY 3578 TTTGTGAGCGCCAGGGTTGGAGTGGGTGGCGGAGGGTGGGGCCAGGCCCTTCATGCAAC 3637
Db 117 ----- 117
QY 3638 GCCCAGCTGTCCGTCGCCGCCAGGGGTGATCTGTCCGCTATGGGCCCGCGTGGGC 3697
Db 118 ----- GlyValPheLeuAlaIatPThrProAlaIatParg 129
QY 3698 GAGCAGAGCGCTTCTCCGTGCCACTTGGCGCAATGGGCTGGGCGCAAGAGTGCCTG 3757
Db 130 GluGlnArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysSerLeu 149

QY 3758 GAGCAGTGGGTGACCGAGGAGCCCGCTCTTTGTCCCGCTTCCGCGCAACAGCCCGGTG 3817
Db 150 GluGlnIrpValThrGluGluAlaAlaCysLeuCysAlaAlaPheAlaAsn- 166
QY 3818 GGTGATGGGCAGAAGGCACAAAGCGGAACTGGAAAGCGGGGACGAGAAAGGCAACC 3877
Db 166 ----- 166
QY 3878 CCTTACCGCATCTCCCAACCCCGAGGAGCCCTTTTCGCCCAACGSCCTCTTGGACAA 3937
Db 167 ----- GlyArgProPheArgProAsnGlyLeuLeuLeuAspLys 180
QY 3938 AGCGGTGAGCAACGTGATCGCTCCCTCACCTGGCGGCGCGCTTTCGAGTACGACGACC 3997
Db 180 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyAspAspPr 200
QY 3998 TCGTCTCTCAGGTCTGAGCTAGCTCAGGAGGACCTGAAGGAGAGTGGGCTTTCT 4057
Db 200 oArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLe 220
QY 4058 GCGGAGGTGGGAGCGAGAGACCGAGGAGTCTCTGACGGCGAGCTCTGAGAGGTGCC 4117
Db 220 uArgGlu- 222
QY 4118 GGGGTGACTGGGGCCTCCGAGGCGAGGATTTGCATAGATGGGTTTGGGAAAGGACAT 4177
Db 222 ----- 222
QY 4178 TCCAGGAGACCCCACTGTAAAGGGCTGGAGGAGGGGACATCTCAGACATGGTGC 4237
Db 222 ----- 222
QY 4238 TGGGAGGTGTGCCCGGTTCAGGGGACACAGGAGAGGCCAAGGACTCTGTACCCCGT 4297
Db 222 ----- 222
QY 4298 CCAGTTGGAGATTTTCGATTTTAGGTTTCTCTCTGGCAAGGAGAGAGGGGTGGAGGTG 4357
Db 222 ----- 222
QY 4358 GCATTGGGAGGAGACTTGGTGAGGTGAGTGAAGACAGGAGGCCCTGGGTCTACCT 4417
Db 222 ----- 222
QY 4418 GGAGTGCCTGGGCCTGAGACTTGTCCAGGTGAACGACAGACACAGGAGGATTGAGAC 4477
Db 222 ----- 222
QY 4478 CCGTTCCTGCTGTGAGTGTGCTGAATGCTGTCCCGTCTCTCTGCAATCCAGCGCT 4537
Db 223 ----- ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe 236
QY 4538 GGCTGGCAAGTCTTACGCTTCCAAAAGGCTTCTGACCCAGCTGGATGAGTGTCTAAC 4597
Db 236 uAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGAGCAGAGATGACCTGGGACCCAGCCAGCCAGCTCTGAGCTGAGTGGGCTTCT 4657
Db 256 rGluHisArgMetThrIrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GGCAAAGAGAGAGGTGAGAGTGGCTGCCACGCTGGGGGCAAGGGTGGTGGTTGAA 4717
Db 276 uAlaGluMetGluLys- 281
QY 4718 GGTCCCAGGAGGAATGAGGGGAGGCTGGGCAAAAGTTGGACCAAGTGCATCACCCGGCA 4777
Db 281 ----- 281
QY 4778 GCCGCATCTGGCTGACAGGTGCAGAATTGGAGGTCAATTTGGGGCTACCCGCTCTATC 4837
Db 281 ----- 281
QY 4838 CCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGGAGGCCCTGAGAGAGCTTCAATGA 4897

```
Db 282 -----AlaLysGlyAsnProGluSerPheAsnAs 292
4898 TGAGAACCTCGCATAGTGGTAACTCTGTCCTGCGGATGCTGACACACCTCGAC 4957
292 pGluAsnLeuArgIleValAlaAspLeuPheSerAlaGlyMetValThrSerTh 312
4958 CAGCTGGCTGGGCTGCTGCTCATGATCTTACACCTGGATGTGACGGTGGAGCCAG 5017
312 rThrLeuAlaTrpGlyLeuLeuMetIleLeuHisProAspValGln----- 328
5018 CTGGGGCCCAAGCAGGACTGAGGAGGAGGTACAGCTGGGGGCCCTGGCTTAGC 5077
328 ----- 328
5078 TGGGACACCGGGCTTCCAGCAGCGCTGGCCAGGCTCTGTAAAGCTTAACCTCTCC 5137
328 ----- 328
5138 AACACAGGAGAGAGAGTGTCCCTGGGTGCTGACCAATTGTGGGGAGCGATGCTGT 5197
328 ----- 328
5198 CCAGTCCGTGTCACACAGAGATCGACGCTGATAGGCGAGGTGGCGCAGCAGATG 5257
329 -Arg--ArgValGlnGlnGluLeuAspValIleGlyGlnValArgArgProGluMet 347
5258 GGTGACCGCTCACATGCCCTACACACTGCGGTGATTACAGAGTGCAGCGCTTTGGG 5317
348 GlyAspGlnAlaHisMetProThrThrAlaValIleHisGluValGlnArgPheGly 367
5318 GACATCATCCCTGAGTGTGACCCATATGACATCCCGTGACATCGAAGTACAGGCTTC 5377
368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
5378 CGCATCCCTAAGTAGGCTGGGGCCCTCTCACCCAGCTCAGCACACGACCTGGTGA 5437
388 ArgIleProLys----- 391
5438 TAGCCCCCAGCATGGCTACTGCCAGGTGGGCCCACTCTAGGAACCTTGGCCACTAGTCT 5497
391 ----- 391
5498 CAATGCCACCACACTGACTGTCCCACTTGGGTGGGGGTCCAGAGTATAGGCGGGCTG 5557
391 ----- 391
5558 GCCTGTCCATCCAGAGCCCCCGTCTAGTGGGGAGACAAACAGGACCTGCCAGATGTTG 5617
391 ----- 391
5618 GAGGACCCAGCGCTGCAGGAGAGGGGCGAGTGTGGTGCCTCTGAGAGGTGTGACTGC 5677
391 ----- 391
5678 GCCCTGCTGTGGGTCCGAGAGGCTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACA 5737
391 ----- 391
5738 GAGTCCAGTGTGTGCCAGGACAGTGTGTCTCCCGGTGTGTGGTGGCAGGGGTCCCAG 5797
391 ----- 391
5798 CATCTAGAGTCCAGTCCCACTCTCACCTGCATCTCTGCCAGGGAACACACTCAT 5857
392 -----Gly-ThrThrLeuIle 396
5858 CACCAACTGTCTATCGGTGCTGAAGATGAGGCGGCTGGGAGAGCCCTTCCGCTTCCA 5917
396 eThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheIle 416
5918 CCCCGAACACTTCTCTGGATGCCAGGGCCACTTTGTGAAGCCGGAGGCTTCTCGCTTT 5977
```

```
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPh 436
5978 CTCAGCAGGTGCTCTGTGGGAGCCCGGCTCCCTGTCTCCCTTCGTTGAGTCTTTCAGGGG 6037
436 e--Ser----- 437
6038 TATACCCAGGAGCCAGGCTCACTGACGCCCTCCCTCCCCACAGGCGCGTGTGATGC 6097
438 -----AlaGlyArgAlaCys 443
6098 CTCGGGAGCCCTCGCCGATGAGCTCTTCTCTTCTTCCCTTCCCTGCTGTCAGCTTT 6157
444 LeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeuLeuGlnHis 463
6158 TTCAGCTTCTCGTGGCGCCGACAGCCCGCCAGCCACTCTCGTGTGTCAGCTTT 6217
464 PheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyValPheAlaPhe 483
6218 CTGTGAGCCCATCCCTTACGAGCTTTGTGTGTGTCGCCCGC 6259
484 LeuValThrProSerProTyrGluLeuCysAlaValProArg 497

RESULT 5
ADB25832
ID ADB25832 standard; protein; 497 AA.
XX
AC ADB25832;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human CYP2D6-related protein #2.
XX
KW human; mutant CYP2D6 gene; drug analysis; drug testing.
XX
OS Homo sapiens.
XX
PN WO2003050282-A1.
XX
PD 19-JUN-2003.
XX
PF 05-DEC-2002; 2002WO-JP012748.
XX
PR 06-DEC-2001; 2001JP-00372548.
XX
PA (TSUR ) TSUMURA & CO.
XX
PI Taniyama M, Ogawa K, Tsuchiya N, Hibino T;
XX
WPI; 2003-505401/47.
XX
DR N-PSDB; ADB25777.
XX
PT Genetic polymorphisms of CYP2D6 gene in human population for analysis of
PT drug effect on individual patients and testing of new drugs.
XX
PS Claim 8; Page 43-46; 75pp; Japanese.
XX
CC The invention comprises mutant forms of the human CYP2D6 gene, containing
CC one or more of the following mutations G125A, C1858T, T2874C and C2875T.
CC The mutant human CYP2D6 genes of the invention are useful for analysing
CC the effect of drugs on individual patients and testing of new drugs. The
CC present amino acid sequence represents a human protein of the invention.
XX
SQ Sequence 497 AA;

Alignment Scores:
Pred. No.: 1.48e-106 Length: 497
Score: 1800.00 Matches: 466
Percent Similarity: 34.14% Conservative: 12
Best Local Similarity: 33.29% Mismatches: 19
Query Match: 9.61% Indels: 903
DB: Gaps: 9

US-09-820-788A-3 (1-10278) x ADB25832 (1-497)
```


Db 222 ----- 222
QY 4298 CCAGTGTGGAGATTTCGATTTTAGTTTCTCTCTGGCAAGAGAGAGGGTGGAGGCTG 4357
Db 222 ----- 222
QY 4358 GCACCTTGGGAGGAGCTTGGTGAGGTCAGTGGTAAGGACAGGAGGCCCTGGGTCTACCT 4417
Db 222 ----- 222
QY 4418 GGAGATGGCTGGGGCCTGAGACTTGTCCAGGTGAACGCAGAGCACAGGAGGATTGAGAC 4477
Db 222 ----- 222
QY 4478 CCGTTCTGTGTGTGTAGTGTGAATGTGTCCCGCTCTCTGCACATCCAGCGCT 4537
Db 223 -----ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe 236
QY 4538 GGTGGCAAGTCTCTACGCTTCCAAAGGTTTCTGACCCAGCTGGATGAGTGTCTAAC 4597
Db 236 uAlaGlyIysValLeuArgPheGlnIysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGAGCACAGGATGACCTGGGACCCAGCCAGCCAGCCAGACCTGACTGAGCCCTTCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GGCAAAGAAGGAAGGTGAGAGTGTGCTGCCAGTGGGGGGCAAGGTTGGTGGTTGAA 4717
Db 276 uAlaGluMetGluLys ----- 281
QY 4718 CGTCCAGGAGGAATGAGGGAGGCTGGGCAAAAGTTGGACAGTGCATCACCCGGCGA 4777
Db 281 ----- 281
QY 4778 GCCGATCTGGGCTGACAGTGCAGAAATTGGAGGTCAATTGGGGGTACCCCGTTCTATC 4837
Db 281 ----- 281
QY 4838 CCCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCAATCA 4897
Db 282 -----AlaIysGlyAsnProGluSerPheAsnAs 292
QY 4898 TGAAACCTCGCATAGTGGTGGTAACCTGTTCCTGGGATGGTGCACCACTCGAC 4957
Db 292 pGluAsnLeuArgIleValAlaAspLeuPheLeuAlaGlyMetValThrThrSerTh 312
QY 4958 CAGCTGGCTGGGGCTCTCTCATGATCTTACACTGGATGTGCAGCTGAGCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGln ----- 328
QY 5018 CTGGGGCCCAAGCAGGAGCTGAGGGAGGAAGGTACAGCTGGGGGCCCTGGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCCGGGGCTTCAGCACAGCGCTGGCCAGGCTCTGTAAAGCCCTTCTCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGGAGGAGAGAGTGTCCCTGGGTGCTGACCCATTGTGGGAGCGCATGTCTGT 5197
Db 328 ----- 328
QY 5198 CCAGTCCGTGTCCAACAGGAGATCGACAGCTGATAGGCGAGTGGCGACACAGAGATG 5257
Db 329 -Arg-ArgValGlnGlnGlnIleAspValIleGlyGlnValArgArgProGluMet 347
QY 5258 GGTGACAGGCTCACATGCCCTACACACTGGCGTGAATTCACGAGTGCAGCGCTTTGGG 5317
Db 348 GlyAspGlnAlaHisMetProTyrThrThrAlaValIleHisGluValGlnArgPheGly 367
QY 5318 GACATCATCCCTGAGTGTGACCCCATATGACATCCGTGACATCGAAGTACAGGCGTTC 5377
|||||:|||||

Db 368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CGCATCCCTTAAGTAGGCTGGCGCCCTCTCTCACCCAGCTCAGCACACGACCTGGTGA 5437
Db 388 ArgIleProLys ----- 391
QY 5438 TAGCCCCAGCATGGCTACTGCGCAGGTGGGCCCACTCTAGGAACCCCTGGCCACCTAGTCT 5497
Db 391 ----- 391
QY 5498 CAATGCCACCACTGACTGTCCCACTTGGGTGGGGGTCCAGAGTATAGGCAGGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTTCATCCAGACCCCGCTCTAGTGGGGAGACAAACAGGACCTCCAGAAATGTTG 5617
Db 391 ----- 391
QY 5618 GAGGACCCAGCCCTGCAGGGAGAGGGGCGAGTGTGGTGCCTCTGAGAGGTGTGACTGC 5677
Db 391 ----- 391
QY 5678 GCCCTGTCTGGGGTCGGAGAGGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACA 5737
Db 391 ----- 391
QY 5738 GAGTCCAGCTGTGCCAGGAGTGTGTGTCTCCCGCTGTGTGGTGGCAGGGTCCACG 5797
Db 391 ----- 391
QY 5798 CATCTAGAGTCCAGTCCCACTCTCACCTGCATCTCTGCCAGGGAAACACACTCAT 5857
Db 392 -----Gly-ThrThrLeuI 396
QY 5858 CACCAACCTGTCTCGTCTGAAGATGAGCGCTCTGGGAGAGCCCTTCCGCTTCCA 5917
Db 396 eThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheH 416
QY 5918 CCCCCAACACTTCTCGGATGCCAGGGCCACTTTGTGAAGCCGGAGCCCTTCTCGCTTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProH 436
QY 5978 CTCAGAGGTGCTGTGGGGAGCCCGCTCTCTGTCCTCCCTTCCGTGGAGTCTTTGAGGG 6037
Db 436 e-Ser ----- 437
QY 6038 TATCACCAGGAGCCAGGCTCACTGACGCCCTCCCTCCACAGGCCCGCGTGCATGC 6097
Db 438 -----AlaGlyArgArgAlaCys 443
QY 6098 CTCGGGAGCCCTGGCCCGCATGGAGTCTTCTCTCTTCCACCTCTCCCTGTCGACGAC 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeuLeuGlnHis 463
QY 6158 TTCAGTCTCTCGTGGCGCCGAGCAGCCCGCCAGCCACTCTCGTGTGTCAGCTTT 6217
Db 464 PheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyValPheAlaPhe 483
QY 6218 CTGTGACCCCATCCCTACGAGCTTTGTGTGTGTGCTGTCGCCCGC 6259
Db 484 LeuValThrProSerProTyrGluLeuCysAlaValProArg 497
RESULT 6
AAW44869
ID AAW44869 standard; protein; 497 AA.
XX
AC AAW44869;
XX
DT 18-SEP-1998 (first entry)
XX
DE Cytochrome P450IId6.
XX Human; cytochrome P450IId6; liver/kidney microsome; LXM-1; antibody;
KW

KW autoimmunohepatitis type II; AIH type II; hepatitis C virus.

XX Hepatitis C virus.

XX Key Location/Qualifiers

FT Misc-difference 65 /note= "encoded by CTT"

FT Misc-difference 66 /note= "encoded by CTT"

FT Misc-difference 72 /note= "encoded by CCG"

FT Misc-difference 72 /note= "encoded by CAN"

FT Misc-difference 287 /note= "encoded by GCG"

FT Misc-difference 374 /note= "encoded by GTG"

FT Misc-difference 388 /note= "encoded by NGC"

XX JP10101698-A.

XX 21-APR-1998.

XX 24-SEP-1996; 96JP-00273055.

XX 24-SEP-1996; 96JP-00273055.

XX (HOKE-) HOKEN KAGAKU KENKYUSHO KK.

XX WPI; 1998-292131/26.

DR N-PSDB; AAV19496.

XX Amino acid sequence for liver/kidney microsomes (LKM) 1 antibody - to

PT human auto-immuno-hepatitis (AIH) type II and hepatitis C.

XX Disclosure; Fig 5; 9pp; Japanese.

XX The present sequence represents the protein of cytochrome P45011d6.

CC Specifically claimed regions of cytochrome P45011d6 are recognised by AIH

CC Type I/II antibody, see AAW44870 and AAW44871 (encoded by

CC AAV19497 and AAV19498). The antigens and their complementing antibodies

CC are used in the detection of auto-immuno-hepatitis type II and hepatitis

CC C

XX Sequence 497 AA;

SQ

Alignment Scores:

Pred. No.: 1.72e-106 Length: 497

Score: 1799.00 Matches: 465

Percent Similarity: 34.14% Conservative: 13

Best Local Similarity: 33.21% Mismatches: 19

Query Match: 9.60% Indels: 903

DB: 2 Gaps: 9

US-09-820-788a-3 (1-10278) x AAW44869 (1-497)

QY 2078 ATGGGGCTAGAGCACTGGTGGCCCTGGCCATGATAGTGGCCATCTTCTGCTCTGCTG 2137

DB 1 MetGlyLeuGluAlaLeuValProLeuAlaMetIleValAlaIlePheLeuLeuVal 20

QY 2138 GACCTGATGACCGGACCAACGCTGGGTGTCACGCTACCGCCAGGTCCCTGCACTG 2197

DB 21 AspLeuMetHisArgGlnArgIrpAlaAlaArgTyrProProGlyProLeuProLeu 40

QY 2198 CCCGGCTGGCAACCTTCTGCTCATGTGGACTTCCAGAACACACCATCTGCTTCACCA 2257

DB 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspG1 60

QY 2258 GGTGAGGAGGAGGTCTCTGGAGGGCGGAGAGGTCTCTGAGGTATGCCCCACACAGCAA 2317

DB 60 n----- 60

QY 2318 CATGGTGTGGTGTAAACCAACACAGCTGGATCAGAAGCCAGGCTGAGAAGGGGAAGCAGG 2377

Db 60 ----- 60

QY 2378 TTTGGGGACGTTCTCTGGGGAAGGACATTTATATCATGTCATGAAGGACTGGATTTCCAA 2437

Db 60 ----- 60

QY 2438 AGGCCAAGGAAGAGTAGGGCAAGGCCCTGGAGGTGGAGCTGGACTTGGCAGTGGGCATGC 2497

Db 60 ----- 60

QY 2498 AAGCCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTTCTGCTGACACCAAGGA 2557

Db 60 ----- 60

QY 2558 AAGCCTTGGGAATGGAAGATGAGTTAGTCTGAGTGCCTTTAAATCAGAAATCGAGG 2617

Db 60 ----- 60

QY 2618 ATGAAGGGGTGTCAGTGACCCGGTTCAAACCTTTTGACACTGTGGTCTCTCGGCCCTCACT 2677

Db 60 ----- 60

QY 2678 GCTCACCGCATGGACCATCATCTGGGAATGGGATGCTAACTGGGGCCTCTCGGCAATTT 2737

Db 60 ----- 60

QY 2738 TGGTGACTCTTTGCAAGGTGATACCTGGGTGAGGCATCCAAACTGAGTTCTCTCCATCAG 2797

Db 60 ----- 60

QY 2798 AAGTGTGACCCCCACCCCTGCCCCAGCATCAGGAGGTGGGTCTCTCTCTCCACCTGC 2857

Db 60 ----- 60

QY 2858 TCACCTCTGTGTAGCCCGGGGTCTGTCCAAAGTTCAAATAGGACTAGGACCTGTAGTCTG 2917

Db 60 ----- 60

QY 2918 GGGTGATCTGGCTTGACAAAGAGGCCCTGACCCCTCTCTGCTGAGTTGGGGCGGCTTCG 2977

Db 61 ----- 66

QY 2978 GGGACGTCTTCAGCTGCAGCTGGCTGGAGCGCGGTGCTGCTCAATGGCTGGCGG 3037

Db 66 lyAspValPheSerLeuGlnLeuAlaIrpThrProValValLeuAsnGlyLeuAla 86

QY 3038 CCGTGGCGAGCGATGCTGACCCGCGCGAGGACACCGCCGACCGCCGCTGGCGCCA 3097

Db 86 laValArgGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProValProI 106

QY 3098 TCTACAGGTCTGGGCTTGGGGCGGCTTCCCAAGGCAAGCGGCTGGGGGACAGAGA 3157

Db 106 leThrGlnIleLeuGlyPheGlyProArgSer-Gln----- 117

QY 3158 CCGGTTTCCGTGGGCCCCGGGTGGACAGTGCACCTAGCCCCCAAGCAGCGCCGACAGGGCG 3217

Db 117 ----- 117

QY 3218 TGGGCTCTGGACGCTGAAACAGATAAAGCCAGAGTGGGTGAGGACAGTGGGCA 3277

Db 117 ----- 117

QY 3278 GGAACCACTGACCGGGGAGGTGCGAGTCTGTGGCTGGAGGGGGCGGGCTACTGC 3337

Db 117 ----- 117

QY 3338 CCAGACCGCCAGAACCCGGTGGGGAGGCTGATGCTCGAAGTGGCGGTGGGGGAC 3397

Db 117 ----- 117

QY 3398 CGCGCTATGCTGGGGCTCAGTGTGGGGGACCGGGCGGATCTTCTTGATGGAAG 3457

Db 117 ----- 117

```
QY 3458 GTGGTCAGGGTGGGACAGAGAGGTTGGGCCAAACCCCGCCCCCAGCAGGAGGCAATG 3517
Db 117 ----- 117
QY 3518 TGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGGCTAGGACTGCCGGAGACC 3577
Db 117 ----- 117
QY 3578 TTGTGAGCGCCAGGGTTGAGTTGGGTGGCGAGGGTGGGGCCCAAGGCCCTTCATGCAAC 3637
Db 117 ----- 117
QY 3638 GCCCAGCTGTCCGTCCGCCCCAGGCGGTGATCCTGTGCGCTATGGGCGCGGTGGCGC 3697
Db 118 -----GlyValPheLeuAlaArgTyrGlyProAlaTrpArg 129
QY 3698 GAGCAGAGCGCTTCTCCGTGTCACCTTCGCAACTTGGGCTGGCGCAAGAGTGCCTG 3757
Db 130 GluGlnArgArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysLysSerLeu 149
QY 3758 GAGCAGTGGGTGACCGAGAGCGCGCTGCTTTGTGCGCTTTCGCCGACCAAGCCGGTG 3817
Db 150 GluGlnTrpValThrGluAlaAlaCysLeuCysAlaAlaPheAlaAsn----- 166
QY 3818 GTGTGAGCGAGAAAGGCAACAAAGCGGAACTGGGAAGCGGGGACCGAGAGGCAACC 3877
Db 166 ----- 166
QY 3878 CCTTACCCGCATCTCCCAACCCAGGAGCGCCCTTCGCCCAACGGGCTCTTGACAA 3937
Db 167 -----His-Ser-----GlyArgProPheArgProAsnGlyLeuLeuAspLys 180
QY 3938 AGCCGTGAGCAACGTGATCGCTCCCTCACTCGGGCGCGCTTCGAGTACGACGACC 3997
Db 180 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPr 200
QY 3998 TCCTCTCTCAGGCTGTGACCTAGCTCAGGAGGACTGAAGGAGAGTTCGGGCTTTCT 4057
Db 200 oArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLe 220
QY 4058 GCGCGAGGTCCGGAGGAGACCGAGGAGTCTCTGCGGGCGGAGCTCTGAGAGGTGCC 4117
Db 220 uArgGlu----- 222
QY 4118 GGGGCTGGACTGGGGCTCCGAAGGCGAGGATTTGCATAGATGGTTCGGAAAGACAT 4177
Db 222 ----- 222
QY 4178 TCAGGAGACCCCACTGTAAAGGCGCTGGAGGAGGGGACATCTCAGACATGGTCG 4237
Db 222 ----- 222
QY 4238 TGGGAGAGGTGTCCCGGGTCAGGGGACCAAGAGAGGCAAGGACTCTGTACCCCGT 4297
Db 222 ----- 222
QY 4298 CCACGTTGGAGATTTCGATTTTAGTTTCTCTCTGGGCAAGGAGAGGGTGGAGGCTG 4357
Db 222 ----- 222
QY 4358 GCACCTTGGGAGGGACTTGGTGAGGTCAGTGTGAAGGACAGGACGCGGCTGTACCT 4417
Db 222 ----- 222
QY 4418 GGAGATGGCTGGGGCTTGACACTTTGTCAGGTGAACGACAGACAGAGGAGGATTGAGAC 4477
Db 222 ----- 222
QY 4478 CCCGTTCTGTGTGTAGTGTGAATGCTGCTCCCGTCTCTCGACATCCAGCGCT 4537
Db 223 -----ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe 236
QY 4538 GGCTGGCAAGGTCCTACGCTTCCAAAAGGCTTTCCTGACCCAGCTGATGAGCTGCTAAC 4597
Db 236 uAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGAGCACAGGATGACCTGGGACCCAGCCAGCCAGCCAGACCTGACCTGAGGCTTCCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GGCAAGAAGAGAGAGTGGCTGCCAGCTGGCGGTGGGGGGCAAGGGTGGTGGGTGA 4717
Db 276 uAlaGluMetGluLys----- 281
QY 4718 CGTCCAGGAGGAATGAGGGAGGCTGGGCAAAAGGTTGGACCAAGTGCATCACC CGCGCA 4777
Db 281 ----- 281
QY 4778 GCGGCATCTGGGCTGACAGGTGCAGAATTGGAGGTCAATTTGGGGGTACCCCGTTCTATC 4837
Db 281 ----- 281
QY 4838 CCTGTGATATCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCAATGA 4897
Db 282 -----AlaLysGlyAsnProGluSerSerPheAsnAs 292
QY 4898 TGAGAACCTGCGCATAGTGTGGGTAACTGTTCTTCTTCCCGGATGGTGACCACTCGAC 4957
Db 292 pGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetValThrThrSerTh 312
QY 4958 CACGCTGGGCTGGGCTCTCTCATGATCTACACCTGGATGTGACGGTGAGGCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuMetIleLeuHisProAspValGln----- 328
QY 5018 CTGGGGCCCAAGGCGAGGACTGAGGGAGGAAGGTTACAGCTGGGGGCCCTCGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCCGGGGCTTCACGACAGGCGTGCGCAGGCTCTGTAAAGCCTAACTTCTCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGGAGGAGGAGAGTGTCCCTGGGTGCTGACCCATTTGTGGGGACGATGCTGT 5197
Db 328 ----- 328
QY 5198 CCAGTCCGTGTCCAAACAGGAGATCGACAGCTGATAGGCGAGGTGGCGACACAGAGATG 5257
Db 329 -Arg-Arg-ArgValGlnGlnIleAspValIleGlyGlnValArgArgProGluMet 347
QY 5258 GGTGACCGGCTCAGATGCCCTACACCTGCGGTGATTTCAGAGGTGACGGCTTTGGG 5317
Db 348 GlyAspGlnAlaHisMetProTyrThrThrAlaValIleHisGluValGlnArgPheGly 367
QY 5318 GACATATCCCTGAGTGTGACCATGACATCCGTGACATCCGATGACATGACAGTACAGGCTTC 5377
Db 368 AspIleValProLeuGlyMetThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CGCATCCCTTAAGTAGGCTGGGGCCCTCTCACCCAGCTCAGCACCCAGCACCTGGTGA 5437
Db 388 ArgIleProLys----- 391
QY 5438 TAGCCCCAGCATGGCTACTGCGAGGTGGGCCCACTCTAGGAACCCCTGGCGCACCTAGCTCT 5497
Db 391 ----- 391
QY 5498 CAATGCCACCACTGACTGTCCCACTTGGGTGGGGGTCCAGAGTATAGGACGGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTTCATCCAGAGACCCCGTCTAGTGGGGAGACAAACAGGACCTGCCAGAAATGTTG 5617
Db 391 ----- 391
QY 5618 GAGGACCCAGCGCTGACGGAGAGGGGGGAGTGTGGGTGCCTCTGAGAGGTGTGACTGC 5677
```


Db 60 n----- 60
QY 2318 CATGGGTGGGTAAACCAAGCTGGATCAGAAAGCCAGGCTGAGAAAGGGAAGCAGG 2377
Db 60----- 60
QY 2378 TTTGGGGAGCTTCCTGGGGAAGACATTTATACATGGCATGAAGACTGGATTTCCAA 2437
Db 60----- 60
QY 2438 AGGCCAAGGAAGTAGGGCAAGGCTGGAGGTGGAGCTGGACTTTGGCAGTGGGCATGC 2497
Db 60----- 60
QY 2498 AAGCCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGTGTACACCAAGAAGGA 2557
Db 60----- 60
QY 2558 AAGGCCTTGGGAATGGAAGATGAGTTAGTCTGTAGTGGCGTTTAAATCAGAAATCGAGG 2617
Db 60----- 60
QY 2618 ATGAAGGGGTGCAGTGACCCGGTTCAAACCTTTTGCACCTGTGGTCTCTCGGSCCTCACT 2677
Db 60----- 60
QY 2678 GCTCACCGGCATGGACCATCATCTGGGAATGGATGCTAACTGGGGCCTCTCGGCAATTT 2737
Db 60----- 60
QY 2738 TGGTGACTCTTCAAGGTGCATACCTGGGTGACGCAATCCAACTGAGTTCTTCATCACAG 2797
Db 60----- 60
QY 2798 AAGGTGACCCCAACCCCTGCCACGATCAGGAGCTGGGTCTCCTCTTCCACCTGC 2857
Db 60----- 60
QY 2858 TCACTCTGTAGTCCCGGGGTCTGTCGAAGGTTCAAATAGGACTAGGACCTGTAGTCTG 2917
Db 60----- 60
QY 2918 GGGTGATCTGTGCTTGACAAGAGCCCTGACCTCTGTCAGTTGGCGCGCTTTCG 2977
Db 61----- 60
QY 2978 GGGACGTGTTACGCTGCAGCTGGCTGGACGCCGGTGGTCTGCTCAATGGGCTGGCG 3037
Db 66 lyaAspValPheSerLeuGlnLeuAlaTrpThrProValValLeuAsnGlyLeuAla 86
QY 3038 CCGTGGCGAGGCGATGGTGACCCGGCGGAGACACGCGCCGACCCGCTGCGCCCA 3097
Db 86 laValargGluAlaLeuValThiGlyGluAspThrAlaAspArgProProValProI 106
QY 3098 TCTACCAGTCTCTGGGCTTGGGCCCGCTTCCCAAGGCAAGCGCGGTGGGAGACAGA 3157
Db 106 leThrGlnIleLeuGlyPheGlyProArgSer-Gln- 117
QY 3158 CCGCGTTTCGTGGGCGCGGTGGACAGTGACCTGACCGGTAGCCCAAGACGCCGACGGCG 3217
Db 117----- 117
QY 3218 TGGGGTCTTGGACGTGAACAGAGATAAAGGCCAGCGAGTGGGCTGAGGACAGTGGGCCA 3277
Db 117----- 117
QY 3278 GGAACCACTGCACGGGGAGGTGGAGTCTCTGTGGCTGGAGGGGGGGCTACTGC 3337
Db 117----- 117
QY 3338 CCAGACCCGCCAGAACCCGGTGGCGAGGCTGATGCGTGAAGTGGCGGTGGCGGGAC 3397
Db 117----- 117

QY 3398 CCGCCTATGCTCGGGCTCAGTGTGGCGGACCGGGCGGGATCTTCTTGTAGTGAAG 3457
Db 117----- 117
QY 3458 GTGGTCAGGGTGGCAGAGACGAGGTGGGGCCAAACCCCGCCAGGAGGAGCAATG 3517
Db 117----- 117
QY 3518 TGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGCTAGGGAATGCGGAGACC 3577
Db 117----- 117
QY 3578 TTGTGAGGCCCAGGGTTGGAGTGGCGGAGGGTGGGCCAAGGCTTTCATGCAAC 3637
Db 117----- 117
QY 3638 GCCCAGTGTCCGTCCCGCCCCCAGGGGTGATCTGTCCGCTATGGGCCCGCGTGGCC 3697
Db 118-----GlyValPheLeuAlaArgTyrGlyProAlaTrpArg 129
QY 3698 GAGCAGAGCGCTTCTCGTGTCCACTTGGCCTGGGCAAGAGTGCCTG 3757
Db 130 GluGlnArgArgPheSerValSerThrLeuArgAsnLeuGlyGlyLysSerLeu 149
QY 3758 GAGCAGTGGGTGACCGAGGAGCCGCTGCTTGTGCGCTTTCGCCGACCAAGCCGCTG 3817
Db 150 GluGlnTrpValThrGluAlaAlaCysLeuCyAlaAlaPheAlaAsn- 166
QY 3818 GGTGATGGCAGAACGGCCAAAGCGGMACTGGGAGCGGGGAGCGGAGAGGCAACC 3877
Db 166----- 166
QY 3878 CCTTACCCGATCTCCCCACCCCGAGAGCGCCCTTTGSCCCCAACGGCTCTTGGACAA 3937
Db 167-----His-Ser- 180
QY 3938 AGCCGTGAGCAACGTGATCGCTTCCCTCACCTGCGGGCGCCGCTTCGAGTACGACACC 3997
Db 180 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPr 200
QY 3998 TCGCTTCTCAGCTGTGGACCTAGCTCAGAGGAGCTGAAGAGGAGTTCGGCTTTCT 4057
Db 200 oArgPheLeuArgLeuLeuAspAlaGlnGluGlyLeuLysGluGluSerGlyPheLe 220
QY 4058 GCGCGAGGTGCGGAGCGAGAGACCGAGGAGTCTCTGACGGGCGAGCTCTCTGAGAGGTGCC 4117
Db 220 uArgGlu- 222
QY 4118 GGGGCTGGAATCTGGGCTCCGAAGGGCAGGATTTGTCATAGATGGGTTTGGGAAGGACAT 4177
Db 222----- 222
QY 4178 TCCAGAGACCCCACTGTGAAGAGGCTGGAGGAGGAGGAGACATCTCAGACATGGTCG 4237
Db 222----- 222
QY 4238 TGGGAGAGGTGTCCCGGGTCAAGGGGACACAGGAGAGGCCAAGGACTCTGTACCCCGT 4297
Db 222----- 222
QY 4298 CCACGTTGGAGATTTTCGATTTTGGTTTCTCCTCTGGCAAGGAGAGAGGGTGGAGGCTG 4357
Db 222----- 222
QY 4358 GCACCTTGGGAGGAGACTTGGTGAAGTCACTGTGTAAAGACAGGACAGGCCCTTGGGTCTACCT 4417
Db 222----- 222
QY 4418 GGAGATGGCTGGGGCTGAGACTTGTCCAGGTGAACGACAGACACAGGAGGATTGAGAC 4477
Db 222----- 222

```
QY 4478 CCCGTTCTGCTGTAGTGTGCTGAATGCTGTCCCGCTCCTCCCTGCACATCCACGCGCT 4537
Db 223 -----ValLeuAsnAlaValProValLeuLeuHisIleProAlaIe 236
QY 4538 GCCTGCAAGGCTCTACGCTTCCAAAGGCTTTCCTGAGCCCAAGCTGGATGCTCTAAC 4597
Db 236 uAlaGlyLyValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuTh 256
QY 4598 TCAGCAGGATGACTGGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GCAAGAGAGAGAGAGTGTGAGAGTGTGCTGCCACGCTGGGGGCAAGGCTGGGTGAA 4717
Db 276 uAlaGluMetGluLys----- 281
QY 4718 CGTCCAGGAGGAATAGGGGAGGCTGGGCAAAAGTTGGACCATCATCACCCGGCGA 4777
Db 281 ----- 281
QY 4778 GCCGCATCTGGGCTGACAGGTGCAGAAATTGGAGGTCAATTGGGGGCTACCCGTTCTATC 4837
Db 281 ----- 281
QY 4838 CCTCGAGTATCCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCAATGA 4897
Db 282 -----AlaLysGlyAsnProGluSerPheAsnAs 292
QY 4898 TGAGAACCTGCGCATAGTGTGGGTAAACCTGTTCTTCGCGGATGGTGACCACTCGAC 4957
Db 292 pGluAsnLeuArgIleValAlaAspLeuPheSerAlaGlyMetValThrSerTh 312
QY 4958 CACGCTGGCTGGGCTCTCTCTCATGATCTCATGCTGATGCTGAGCTGAGCGTGGAGCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuMetIleLeuHisProAspValGln----- 328
QY 5018 CTGGGGCCCAAGCGAGGACTGAGGGAGGAGGTACAGCTGGGGGCCCTGGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCGGGGCTTCAGCACAGCGGTGGCCAGGCTCTGTAAAGCTTAACCTCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGGAGGAGAGAGTGTCCCTGGGTGCTGACCCATTGTGGGACGCGATGCTGT 5197
Db 328 ----- 328
QY 5198 CCAGTCCGTGTCCAAAGAGAGATCGACGACGTGATAGGACAGTGTGGCGACACAGAGATG 5257
Db 329 -Arg--ArgValGlnGlnGluLeuAspValIleGlyGlnValArgProGluMet 347
QY 5258 GGTGACCAAGGCTCACATGCCCTACACCACTGCGTGTATTCACGAGGTGCGCGCTTTGGG 5317
Db 348 GlyAspGlnAlaHisMetProLysThrThrAlaValIleHisGluValGlnArgPheGly 367
QY 5318 GACATCATCCCCCTGAGTGTGACCATATGACATCCCTGACATCGAAGTACAGGCTTC 5377
Db 368 AspIleValProLeuGlyMetThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CACATCCCTAAGGTAGGCTGGCGCCCTCTCACCCAGCTCAGCACACCACTGGTGA 5437
Db 388 ArgIleProLys----- 391
QY 5438 TAGCCCCCAGCATGGCTACTGCGAGGTGGGCCCACTCTAGGAACCTGTGGCCACTAGTCCT 5497
Db 391 ----- 391
QY 5498 CAATGCCACCACTGACTGTCCCACTTGGGTGGGGGTCCAGAGTATAGGCAGGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTCCATCCAGAGCCCCCGTCTAGTGGGGAGACAAACAGGACCTGCCAGAAATTGTG 5617
```

```
Db 391 ----- 391
QY 5618 GAGGACCAGCGCCTGACAGGAGAGAGGGGAGTGTGGTGCCTCTGAGAGGTGTGACTGC 5677
Db 391 ----- 391
QY 5678 GCCCTGCTGTGGGTTCGAGAGAGGCTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACA 5737
Db 391 ----- 391
QY 5738 CAGTCCAGCTGTGTGCCAGGAGTGTGTGTCCCGCTGTGTGTGGTGGCAGGGGTCCAG 5797
Db 391 ----- 391
QY 5798 CATCTAGATCCAGTCCCACTCTCATCCCTGCATCTCTCCCGCAGGAAACGACATCAT 5857
Db 392 -----Gly-ThrThrLeuIle 396
QY 5858 CACCAACCTGTATCGGTGCTGAAGATGAGGCGCTCTGGGAGAGCCCTTCCGCTTCCA 5917
Db 396 eThrAsnLeuSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHi 416
QY 5918 CCCCAGAACACTTCTGATGCCAGGGCCACTTTGTGAAGCCGGAGGCTTCTCCCTGCTTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPh 436
QY 5978 CTCAGCAGGTGCTGTGGGGAGCCCGCTCCTCTCCCTTCCCTGGAGTCTTCGAGGGG 6037
Db 436 e--Ser----- 437
QY 6038 TATCACCAGGAGCCAGGCTCACTGACGCCCTCCCTCCCGCAGGCGCGCTGCATGC 6097
Db 438 -----AlaGlyArgAlaCys 443
QY 6098 CTCGGGAGAGCCCTGGCCCGCATGAGCTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeuLeuGlnHis 463
QY 6158 TTCAGCTTCTCCGTGGCGCCGCGAGCAGCCCGCGCCAGCCACTCTCGTGTGCTGAGCTTT 6217
Db 464 PheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyValPheAlaPhe 483
QY 6218 CTGTGTACCCCTCCCTACGAGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6259
Db 484 LeuValThrProSerProLysGluLeuCysAlaValProArg 497

RESULT 8
AAR72377
ID AAR72377 standard; protein; 497 AA.
XX
AC AAR72377;
XX
DT 25-MAR-2003 (revised)
DT 15-NOV-1995 (first entry)
XX
DE Human auxillary cytochrome P450 species 2D6 variant 2 protein.
XX
KW Human cytochrome P450; amplification; PCR; primer; expression vector;
KW Yeast NADPH-P450 reductase; safety; fusion protein; metabolite;
KW carcinogen; mutagen; liver metabolism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 296 /note= "Cys to Arg variation"
FT FT
FT Misc-difference 486 /note= "Thr to Ser variation"
FT FT
XX EP644267-A2.
XX
XX 22-MAR-1995.
```



```

XX 20-JUL-1994; 94EP-00111298.
XX PF XX
XX 20-JUL-1993; 93JP-00201120.
XX PR XX
XX 21-JUL-1993; 93JP-00180246.
XX PR XX
XX 30-JUL-1993; 93JP-00208279.
XX PR XX
XX (SUMO) SUMITOMO CHEM CO LTD.
XX PA XX
XX PI XX
XX Hayashi K, Sakaki T, Yabusaki Y, Komai K, Kaneko H, Nakatsuka I;
XX XX
XX WPI; 1995-116991/16.
XX DR XX
XX N-ESDB; AAQ87731.
XX DR XX
XX Evaluation of safety of a chemical cpd. - using recombinant yeast
XX PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase.
XX PR
XX Example; Page 91-93; 124pp; English.
XX PS
XX CC The amino acid sequence of the human auxiliary cytochrome P450 species
XX CC 2D6 variant 2. This variant contains variations at residues 296: Cys to
XX CC Arg and 486: Thr to Ser, caused by variations at bases 886: T to C and
XX CC 1457: C to G in the DNA sequence. The cDNA was amplified by PCR using the
XX CC primers AAQ87763-6. The product was cloned into the yeast expression
XX CC vectors pAAH5N or pAHRr to produce the vectors p2D6 variant 2 for co-
XX CC expression of the cytochrome P450 alone or p2D6R variant 2 for co-
XX CC expression with the yeast NADPH-P450 reductase. The vectors are used in a
XX CC method for evaluating the safety of a chemical compound by reacting the
XX CC chemical compound with recombinantly produced human cytochrome P450
XX CC molecular species 1A2 (AAQ87714), 2C9 (AAQ87715), 2E1 (AAQ87716), or 3A4
XX CC (AAQ87718), or their auxiliary species and variants (AAQ87718-32), and
XX CC yeast NADPH-P450 reductase, either as a fused protein or in cell
XX CC extracts, and analysing the resulting metabolite to assess the safety of
XX CC the chemical compound. The method is useful for determining whether the
XX CC chemical compound, or its metabolite, will be converted into a
XX CC carcinogenic or mutagenic form through metabolism in the liver. (Updated
XX CC on 25-MAR-2003 to correct PN field.)
XX CC
XX SQ Sequence 497 AA;

Alignment Scores:
Pred. No.: 1.99e-106 Length: 497
Score: 1798.00 Matches: 465
Percent Similarity: 34.14% Conservative: 13
Best Local Similarity: 33.21% Mismatches: 19
Query Match: 9.60% Indels: 903
DB: 2 Gaps: 9

US-09-820-788A-3 (1-10278) x AAR72377 (1-497)

QY 2078 ATGGGGCTAGAACGACCTGGTCCCTCGGCATGATAGTGGCCATCTTCCTGCTCTGCTGTG 2137
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MeGlyLeuGluAlaLeuValProLeuAlaValIleValIlePheLeuLeuVal 20

QY 2138 GACCTGATGACCGGCACCAACGCTGGGGCTGCACGCTACCCGCGAGGTCCCTGCCACTG 2197
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 AspLeuMethIleArgArgGlnArgTriaAlaAlaArgTyrProGlyProLeuProLeu 40

QY 2198 CCGGGCTGGGCAACCTTGCTGCGATGTGGAGCTTCACAGAACACCATCTGCTTCGACCA 2257
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41 ProGlyLeuGlyAsn-LeuLeuHisValAppPheGlnAsnThrProTyrCysPheAspG1 60

QY 2258 GGTGAGGAGGAGGTCTCTGGAGGGCGCAGAGGTCTTGAGGATGCCCCACCACGACAA 2317
Db | n----- 60
Db 60 n----- 60

QY 2318 CATGGTGGTGGGTTAAACCAACGAGCTGGATCAGAAGCCAGGCTGAGAGGGGAAGCAGG 2377
Db ----- 60
Db 60 ----- 60

QY 2378 TTTGGGGGACGTTCTCTGGGGAAGGACATTTATCATGGCATGAAGGACTGTGATTTTCCAA 2437
Db ----- 60
Db 60 ----- 60

```

QY 3518 TGGGTGAGCAAGAGTGGGCCCTGTGTCCAGCTGACCGGCTAGGGACTGCGGAGACC 3577
Db 117 ----- 117
QY 3578 TTGTGGAGCGCCAGGGTTGGAGTGGGTGGCGAGGGTGGGGCAAGGCCTTCATGGCAAC 3637
Db 117 ----- 117
QY 3638 GCCCAGCTGTCGTCGCCGCCCGCCAGGGGTGATCCTGTGCGGTATGGCCCGCGTGGCGC 3697
Db 118 -----GlyValPheLeuAlaArgTyrGlyProAlaTrpArg 129
QY 3698 GAGCAGAGCGCTTCTCCGTGTCCACCTTGCCAACTTGGGCTGGGCAAGAGTCGCTG 3757
Db 130 GluGlnArgArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysSerLeu 149
QY 3758 GAGCAGTGGGTGACGAGAGCGCCCTGCTTGTGCGCTTCCGCCAGCAAGCCGCTG 3817
Db 150 GluGlnTrpValThrGluGluAlaCysLeuCysAlaAlaPheAlaAsn----- 166
QY 3818 GGTGATGGGCAGAGGCGCAAGCGGNACTGGGAAGCGCGGACGAGNAGGCAACC 3877
Db 166 ----- 166
QY 3878 CCTTACCGCATCTCCCGACCCCGCAGGACGCCCTTTGCCCCAACCGCCTCTGGACAA 3937
Db 167 -----His-Ser-----GlyArgProPheArgProAsnGlyLeuLeuAspLys 180
QY 3938 AGCCTGAGCAACGTGATGCGTCCCTCACCTGCGGGCGCGCTTCGAGTAGCAGCACCC 3997
Db 180 sAlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPr 200
QY 3998 TGCCTTCTCAGGCTGCTGGACCTAGCTCAGAGGGGACTGAAGGAGGAGTGGGCTTCT 4057
Db 200 oArgPheLeuArgLeuLeuAspLeuAlaGlnGlyLeuLysGluGluSerGlyPheLe 220
QY 4058 GCGCAGGTGCGGAGCGAGAGACCGAGGAGTCTCTGCGGGCGAGCTCTGTAGAGGTGCC 4117
Db 220 uArgGlu----- 222
QY 4118 GGGGTGAGTGGGGCCTCCGAAGGCGAGGATTTGCATAGATGGGTTTGGGAAAGACAT 4177
Db 222 ----- 222
QY 4178 TCCAGGAGCCCACTGTAAAGAGGCGCTGGAGGAGGGGACATCTCAGACATGGTGC 4237
Db 222 ----- 222
QY 4238 TGGGAGAGGTGTGCCCGGTGAGGGGCGCACGAGGAGGCGCAAGGACTCTGTACCCCGT 4297
Db 222 ----- 222
QY 4298 CCACGTTGGAGATTTTCGATTTTTCCTCTCTGGCAAGGAGAGGGTGGAGGCTG 4357
Db 222 ----- 222
QY 4358 GCACTTGGGAGGAGACTTGGTGAGTCACTGTGTAGGACAGCGAGGCGCTGGTCTACCT 4417
Db 222 ----- 222
QY 4418 GGAGATGGTGGGGCTGAGACTTGTCCAGGTGAACGCGAGAGCACAGGAGGATTGAGAC 4477
Db 222 ----- 222
QY 4478 CCGTTCTGTCTGGTGTAGGTGCTGAATGCTGTCTCCCGTCTCTGTCACATCCCGCGCT 4537
Db 223 -----ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe 236
QY 4538 GGCTGGCAAGTCTCAGCTTCCAAAGCTTCTGACCCAGCTGGATGAGCTGCTAAC 4597
Db 236 uAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGAGCACAGGATGACCTGGGACCCAGCGCCAGCCAGGACCTGTGACTGAGGCGCTTCTCT 4657

Db rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GGCMAAGAAAGAGAGGTGAGAGTGGCTGCCACCGTGGGGGGCAAGGGTGGTGGTTGAA 4717
Db 276 uAlaGluMetGluLys----- 281
QY 4718 CGTCCAGGAGGAATGAGGGGAGGCTGGGCAAAAGTTGGACCACTGATCACCACCGCGCA 4777
Db 281 ----- 281
QY 4778 GCGCATCTGGGCTGACAGGTGCAGATTGGAGTTCATTTGGGGCTACCCCGTTCTATC 4837
Db 281 ----- 281
QY 4838 CCCTGAGTATCCTCTCGGCCCTGTCTCAGGCAAGGGAGCCCTGAGAGCAGTTCATGTA 4897
Db 282 -----AlaLysGlyAsnProGluSerSerPheAsnAs 292
QY 4898 TGAGAACCTGCGCATAGTGGGTAACTCTTCTTCCCGGATGTGACACCTCGAC 4957
Db 292 pGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetValThrThrSerTh 312
QY 4958 CACGTCGCTCGGGCCCTCTCTCTCATCATCTTACACCTGATGTGACGCTGAGCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuMetIleLeuHisProAspValGln----- 328
QY 5018 CTGGGGCCCAAGGAGGAGCTGAGGAGGAAGGTACAGCTGGGGGCCCCCTGGGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCCGGGCTTCCAGCACAGGCGTGGCCAGGCTCTGTAAAGCCTTAATCTCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGGAGGAGAGAGTGTCCCTCGGTGCTGACCCATTGTGGGACGATGTCTGT 5197
Db 328 ----- 328
QY 5198 CCAGTCCGTGTCCAAAGGAGATCGACGCTGATAGGGCAGGTGCGGCGACGAGATG 5257
Db 329 -Arg--ArgValGlnGlnGluIleAspAspValIleGlyGlnValArgArgProGluMet 347
QY 5258 GGTGACAGGCTCACATGCGCTTACACCACTGCGGTGATTACAGAGTGCAGCGCTTGGG 5317
Db 348 GlyAspGlnAlaHisMetProTyrThrAlaValIleHisGluValGlnArgPheGly 367
QY 5318 GACATCATCCCTGAGTGCACCATATGACATCCCGTGACATCGAAGTACAGGGCTTC 5377
Db 368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CGCATCCCTAAGGTAGGCTGGCGCTCTCTCAACCCAGCTCAGCACGAGCCTGGTGA 5437
Db 388 ArgIleProLys----- 391
QY 5438 TAGCCCCAGCATGGTACTGCCAGGTGGGCCCACTCTAGGAACCTTGCGCCACCTAGTCT 5497
Db 391 ----- 391
QY 5498 CAATGCCACACACTGACTGTCTCCCACTTGGGTGGGGGTCCAGAGTATAGSCAGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTCCATCCAGAGCCCCGTCTAGTGGGGGAGACAAACAGGACCTGCCAGAAATGTTG 5617
Db 391 ----- 391
QY 5618 GAGGACCCAGCGCTGCAGGAGAGGGGCGAGTGTGGTGCTCTGAGAGGTGTGACTGC 5677
Db 391 ----- 391
QY 5678 GCCCTGTGTGGGTGCGAGAGGGTACTGTGGAGCTTCTCGGGGCGGAGGACTAGTTGACA 5737

Db 60 ----- 60
QY 2738 TGGTGACTCTTGCAGAGTCAATACCTGGGTGACGATCCAAACTGAGTTCCTCCATCACAG 2797
Db 60 ----- 60
QY 2798 AAGGTGTGACCCCAACCCCTGCCCAAGATCAGGAGGTGGGTCTCTCTCCATCCACCTGC 2857
Db 60 ----- 60
QY 2858 TCACTCTGTAGTCCCGGGTCTGTCAAAGTTCAAATAGGACTAGGACCTGTAGTCTG 2917
Db 60 ----- 60
QY 2918 GGGTGATCCTGGCTTTGACAAGAGGCCCTGAACCTCCCTCTGCGATTGGCGCCGCTTCG 2977
Db 61 ----- 66
QY 2978 GGGACGTGTTCACTGACCTGGCTGGACCGCCGGTGGTCTGCTCAATGGGCTGGCGG 3037
Db 66 lYAspValPheSerLeuGlnLeuAlaTrpThrProValValLeuAsnGlyLeuAlaA 86
QY 3038 CGGTGCGGAGCGGATGTGACCGCGGAGGACACCGCCGACCGCCGCTGGCCCA 3097
Db 86 laValargGluAlaLeuValThrHisGlyGluaspThrAlaaspargProProValProI 106
QY 3098 TCTACAGGTCTCTGGCTTCGGCGCGCTTCCAAAGCAAGCGCGGTGGGGGACAGAGA 3157
Db 106 leThrGlnileLeuGlyPheGlyProArgser-Gln- 117
QY 3158 CGCGCTTCCGTGGCGCCCGGTGGACAGTGCACCTAGCCCAAGCAGCGCCGACAGGGCG 3217
Db 117 ----- 117
QY 3218 TGGGTCTCTGGAGTGAACAGAGATAAGGACGAGTGGGCTGAGGACAGTGGGCA 3277
Db 117 ----- 117
QY 3278 GGAACACCACTGCACGGGGAGGTGCGAGTCTGTGGGTGGGAGGGGGCGGGCTACTGC 3337
Db 117 ----- 117
QY 3338 CCAGACCCGCAAGACCCGCTGGGGAGGCTGATGCTGAGTGGCGGTGGCGGGGAC 3397
Db 117 ----- 117
QY 3398 CGCGCTATGCTGCGGGCTCAGTGTGGCGGACGCGGGGATCTTCTTGAGTGGAAAG 3457
Db 117 ----- 117
QY 3458 GTGGTCAAGGTGGGAGAGACAGAGTGGGGCCAAACCCCGCCAGGCGGGGAGCAATG 3517
Db 117 ----- 117
QY 3518 TGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGACCGGGCTAGGACTGGGGAGACC 3577
Db 117 ----- 117
QY 3578 TTGTGGAGCGCCAGGCTTGGAGTGGGTGGCGAGGTGGGGCCAAAGCCCTTCATGGCAAC 3637
Db 117 ----- 117
QY 3638 GCCACAGTGTCCGTCGCCGCCAGGGGTGATCCTGTGCGCTATGGCGCCGCTGGCGC 3697
Db 118 ----- 129
QY 3698 GAGCAGAGCGCTTCTCGTGTCCACTTGGCCCACTTGGGCTGGGCAAGAGTTCGCTG 3757
Db 130 GluGlnArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysLysSerLeu 149
QY 3758 GAGCAGTGGGTGACCGAGGAGCGCTTGTGGCGCTTCTGCGGCTTCGCGCAGCAAGCGGTG 3817

Db 150 GluGlnTrpValThrGluGluAlaAlaCysLeuCysAlaAlaPheAlaAsn- 166
QY 3818 GGTGATGGCAGAGGGCACAAACGGGAACTGGGAAGCGGGGACGGAGAAGCAACC 3877
Db 166 ----- 166
QY 3878 CTTTACCGGATCTCCCAACCCCGAGGACGCCCTTTTCGCCCAACCGCTCTTTGGACAA 3937
Db 167 ----- 180
QY 3938 AGCGTGAGCAACGTGATCGCTCCCTACCTCGGGCGCGCTTCGAGTACGACGACCC 3997
Db 180 sAlaValSerAsnValIleAlaSerLeuThrCysGlyArgPheGluTrpAspPr 200
QY 3998 TCGCTTCTCAGGCTGTGGACCTAGCTCAGGAGGACTGAAGAGGAGTGGGCTTTCT 4057
Db 200 oArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLe 220
QY 4058 CGCGAGGTGGGAGCGAGACCGAGAGTCTCTGCAGGGCGAGCTCTCAGAGGTGCC 4117
Db 220 uArgGlu- 222
QY 4118 GGGCTGACTGGGGCTCCGAAGGCGAGGATTTGCATAGATGGGTTTGGAAAGGACAT 4177
Db 222 ----- 222
QY 4178 TCCAGGAGACCCACTGTAGAAGGGCTGGAGGAGGGGACATCTCAGACATGGTCG 4237
Db 222 ----- 222
QY 4238 TGGGAGGTGTGCCCGGTTCAGGGGCAACGAGAGGCCAAGAGTCTCTGACCCCGT 4297
Db 222 ----- 222
QY 4298 CCACGTTGGAGATTTCGATTTTAGGTTTCTCTCTGGCAAGGAGAGAGGTGGAGGTG 4357
Db 222 ----- 222
QY 4358 GCACCTGGGAGGACTTGGTGAGTCACTGAGTAAAGNACAGGAGGCGCTGGGTCTACT 4417
Db 222 ----- 222
QY 4418 GGAGATGCTGGGCTGAGACTTGTCCAGGTGAACGAGAGCACAGAGGATGAGAC 4477
Db 222 ----- 222
QY 4478 CCGGTTCTGTCTGTGTAGTGTCAATGTCTCCCGTCTCTCTGCACTCCAGCGCT 4537
Db 223 ----- 236
QY 4538 GGCTGGCAGGTCTTACGCTTCCAAAGGCTTTCTGACCCAGCTGGATGAGTCTTAAC 4597
Db 236 uAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGAGCACAGGATGACCTGGGACCCAGCCAGCCAGCTGCTGAGCTGAGGCGCTTCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GSCAAAGAGAGAGGTGAGAGTGGCTGCCACCGTGGGGGCAAGGTTGGGTGAA 4717
Db 276 uAlaGluMetGluLys- 281
QY 4718 CGTCCAGGAGGATGAGGGGAGGCTGGGCAAAAGTTGGACAGTGCATCACCAGCGCA 4777
Db 281 ----- 281
QY 4778 GCCGATCTGGCTCAGAGGTGCAGATTGGAGTTCATTGGGGCTACCCGCTTCTATC 4837
Db 281 ----- 281
QY 4838 CCTGAGTATCTCTCGGCCCTGTCTAGGCCAAGGGAGCCCTGAGACAGCTTCAATGA 4897
Db 282 ----- 292
AlaLysGlyAsnProGluSerPheAsnAs

QY 4898 TGAGAACCTCGCATAGTGTGGGTAACTGTTCTCTTGGCGGATGTGACCACTCGAC 4957
Db |||||
292 pGluAsnLeuArgileValValAlaAspLeuPheSerAlaGlyMetValThrThrSerTh 312
QY 4958 CAGCTGGCTGGCGCTCTCTGCTCATGATCTTACACCTGATGTGACGGTGCAGCCAG 5017
Db |||||
312 rThrLeuAlaIrrpGlyLeuLeuLeuMetIleLeuHisProAspValGln----- 328
QY 5018 CTGGGGCCCAAGGCAGGAGCTGAGGGAGAGGGGTACAGCTGGGGGCCCTGGGCTTAGC 5077
Db ----- 328
QY 5078 TGGGACACCCGGGGCTTCCAGCACAGGCGTGGCCAGGCTCTGTAAAGCTTAATCTCTCC 5137
Db ----- 328
QY 5138 AACACAGGAGGAGAGAGTGTCCCTGGGTGTGCTGACCCATTGTGGGGAGCGATGTCTGT 5197
Db ----- 328
QY 5198 CCAGTCCGTCTCAACAGAGATCGACAGCTGATAGGGCAGGTGGCGGACACAGAGATG 5257
Db ::|||
329 -Arg-ArgValGlnGlnGlnIleAspAspValIleGlyGlnValArgArgProGluMet 347
QY 5258 GTGACAGGCTCACATGCCCTTACACACTGCCGTGATTACAGAGTGCAGCGCTTTGGG 5317
Db |||||
348 GlyAspGlnAlaHisMetProTyrThrAlaValIleHisGluValGlnArgPheGly 367
QY 5318 GACATCATCCCTGAGTGTGACCCATATACATCCGTGACATCAAGTACAGGCTTC 5377
Db ::|||
368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CGCATCCCTAAGTAGGCTGGCGCTCTCTCACCCAGCTCAGCAGCAGCAGCTGTGTA 5437
Db |||||
388 ArgIleProlys----- 391
QY 5438 TAGCCCCAGCATGGTACTGCCAGGTGGGCCACTCTAGGAACCTTGGCCACTAGTCTCT 5497
Db ----- 391
QY 5498 CAATGCCACCACTGACTGTCCCCACTTGGTGGGGGTCCAGAGTATAGGCGGGCTG 5557
Db ----- 391
QY 5558 GCCTGTCCATCCAGAGCCCCGTCTAGTGGGGAGACAAACAGGACCTGCCAGAAATTG 5617
Db ----- 391
QY 5618 GAGGACCCAGCGCTGCAGGAGAGGGGGCAGTGTGGGTGCTCTGAGAGGTGTGACTGC 5677
Db ----- 391
QY 5678 GCCTGTGTGGGGTCGGAGAGGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACA 5737
Db ----- 391
QY 5738 GAGTCCAGCTGTGTGCCAGCAGTGTGTCCCGCTGTGTTGGTGGCAGGGGTCCCAG 5797
Db ----- 391
QY 5798 CATCTTAGATCCAGTCCCACACTCTCACCCCTGCATCTCTGCCAGGGAACAGACACTCAT 5857
Db |||||
392 -----Gly-ThrThrLeuIle 396
QY 5858 CACCAACTGTCTATCGTGTGATGAAGATGAGCGCTGTGGGAGAACCCCTTCGCTTCCA 5917
Db |||||
396 eThrAsnLeuSerSerValLeuIlysAspGluAlaValTrpGluIysProPheArgPheHis 416
QY 5918 CCCCAGACATCTCTGATGCCAGGCCACTTTGTGAAGCCGAGGCTTCTGCTGCTTT 5977
Db |||||
416 sProGluHisPheLeuAspAlaGlnGlyHisPheValIysProGluAlaPheLeuProPh 436

QY 5978 CTCAGCAGTGTGCTGTGGGAGCCCGGCTCCCTGTCCCTTCCGTGGAGTCTTTCAGGGG 6037
Db |||||
436 e--Ser----- 437
QY 6038 TATCACCCAGGAGCAGGCTCACTGACGCCCTCCCTCCCCACAGCGCCGCTGCATGC 6097
Db -----AlaGlyArgArgAlaCys 443
QY 6098 CTCGGGAGCCCTGGCCGATGAGTCTTCTCTTTCACCTCCCTGCTGCAGCAC 6157
Db |||||
444 LeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeuLeuGlnHis 463
QY 6158 TTCAGCTTCTCGTGGCCCGGACAGCCCGGCGCAGCAGCTCTCTGTGTGTCAGCTTT 6217
Db |||||
464 PheSerPheSerValProThrGlyGlnProArgProSerHisGlyValPheAlaPhe 483
QY 6218 CTGGTGACCCCTCCCTACGAGCTTGTGTGTCGCCCGC 6259
Db |||||
484 LeuValSerProSerProTyrGluLeuCysAlaValProArg 497

RESULT 10
AAR81462
ID AAR81462 standard; protein; 497 AA.
XX
AC AAR81462;
XX
DT 01-AUG-1996 (first entry)
XX
DE Human derived cytochrome P4502D6.
XX
KW Human derived cytochrome; P4502D6; commercial cDNA library; yeast;
KW transfection; recombinant production; expression vector; mammal;
KW immunisation; sensitisation; antibody; determination; detection;
KW non-cross reactive.
XX
OS Homo sapiens.
XX
PN JP08027199-A.
XX
PD 30-JAN-1996.
XX
PF 15-JUL-1994; 94JP-00164186.
XX
PR 15-JUL-1994; 94JP-00164186.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
DR WPI; 1996-136340/14.
DR N-PSDB; AAT17388.
XX
PT Antibody recognising human derived cytochrome P4502D6 - allows specific
PT detection of cytochrome P450 species in humans.
XX
PS Example 1; Page 11-13; 13pp; Japanese.
XX
CC The present sequence is the human derived cytochrome (HDC) P4502D6, which
CC was obtd. from a commercial cDNA library. Yeast were transfected with an
CC expression vector contg. the HDC cDNA, cultured and then disrupted to
CC give a microsomal fraction. The HDC was purified from the fraction, and
CC used to immunise and sensitise a mammal. Blood was drawn from the mammal,
CC and an anti-HDC antibody isolated. The antibody obtd. recognises HDC
CC P4502D6, partic. at a serum dilution rate of 1:10000, and is
CC substantially without cross reaction to other HDC P450 spp.
XX
SQ Sequence 497 AA;
Alignment Scores:
Pred. No.: 1,99e-106 Length: 497
Score: 1798.00 Matches: 465
Percent Similarity: 34.14% Conservative: 13
Best Local Similarity: 33.21% Mismatches: 903
Query Match: 9.60% Indels: 9
DB: 2 Gaps: 9

US-09-820-788A-3 (1-10278) x AAR81462 (1-497)

QY 2078 ATGGGGCTAGAAAGCATGCTGGTGGCCCTGGCCATGATAGTGGCCATCTTCTGTCTCTGTG 2137
Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
QY 2138 GACCTGATGACCGGCACCAAGCTGGCTGACAGCTACCGCCAGGTCCCTGCACTG 2197
Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTrpProGlyProLeuProLeu 40
QY 2198 CCCGGCTGGGCAACCTTCTGCTCATGTGACTTCCAGAACACACCATACTGTCTGACCA 2257
Db 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspG1 60
QY 2258 GGTGAGGAGGAGGTCTTGGAGGGCGGAGAGGTCTGAGGATGCCCCACACAGCAAA 2317
Db 60 n----- 60
QY 2318 CATGGGTGGTGGTTAAACCAACAGCTGGATCAGAAGCCAGGCTGAGAAGGGAGAGCAGG 2377
Db 60----- 60
QY 2378 TTTGGGGACGTTCTCTGGGAAGGACATTTATACATGGCATGAAGGACTGGATTTTCCAA 2437
Db 60----- 60
QY 2438 AGGCCAAGAGAGTAGGCAAGGCCCTGGAGGTGGAGCTGGAGCTGGAGTGGGCATGC 2497
Db 60----- 60
QY 2498 AAGCCCATTTGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACACCAGAGA 2557
Db 60----- 60
QY 2558 AAGGCCTTGGGAATGGAAGTAGTGTAGTCTGAGTGGCGTTTAAATCACGAATTCGAGG 2617
Db 60----- 60
QY 2618 ATGAAGGGGGTGCAGTAGACCCCGTTCAACCTTTTGCACGTGTGGTCTCGGGCCCTACT 2677
Db 60----- 60
QY 2678 GCTACCGGCATGGACCATCATCTGGGAATGGATGGATGCTAACTGGGGCCTCTCGGCAATT 2737
Db 60----- 60
QY 2738 TGGTGACTCTTCAAGGTATACCTGGGTGACGATCCAAACTGAGTTCTCTCCATCACAG 2797
Db 60----- 60
QY 2798 AAGGTGTGACCCCAACCCCTGCCCCAGATCAGGAGGTGGGTCTCTCTCTTCCACCTGC 2857
Db 60----- 60
QY 2858 TCACCTCTGTAGTACCCCGGGGTCTGCCAGGTTCCTCCAGTTCCTGAGTCTG 2917
Db 60----- 60
QY 2918 GGGTGATCTGTGCTTTGACAAGAGGCCCTGACCCCTCTGTCAGTGTGGCGCCGCTTCG 2977
Db 61----- 66
QY 2978 GGGACGTGTTCAGCTGCAGCTGGCTGGACCGCGGTGGTCTGTCTCAATGGGCTGGCGG 3037
Db 66 lyAspValPheSerLeuGlnLeuAlaTrpThrProValValLeuAsnGlyLeuAla 86
QY 3038 CCGTCGGGAGCGATGTGACCCCGCGGAGGACACCGCCGACCGCCGCTGCGCCCA 3097
Db 86 laValArgGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProProValProI 106
QY 3098 TCTACAGGTCTCTGGGCTTGGGGCGCGTTTCCCAAGCAAGCGCGGTGGGGGACAGAGA 3157
Db 117----- 222

Db 106 leThrGlnIleLeuGlyPheGlyProArgSer-Gln----- 117
QY 3158 CCGGTTTCCGTGGGCCCCGGGTGGACAGTACCGTAGCCCAACAGAGCCCGACAGGGCG 3217
Db 117----- 117
QY 3218 TGGGTCCTGACCGTGAACACAGATAAAGCCAGAGTGGGCTGAGGACAGTGGGCCA 3277
Db 117----- 117
QY 3278 GGAACACCACTGCACGGGGAGGTGCGAGTCTGTGGCTGGGAGGGGCGGGGCTACTGC 3337
Db 117----- 117
QY 3338 CCAGACCCGCAGAAAGCCCGTGGCGGAGGCTGATCGTCAAGTGGCGGTGGCGGGAC 3397
Db 117----- 117
QY 3398 CGCGCTATGCTCGGGCTCAGTGTGGCGGACGGCGGATCTTCTTGTGAGTGAAG 3457
Db 117----- 117
QY 3458 GTGTCAGGGTGGCAGAGACAGGTGGGCGCAAAACCCGCCAGGAGGGAGCAATG 3517
Db 117----- 117
QY 3518 TGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGGCTAGGACTCGGGAGACC 3577
Db 117----- 117
QY 3578 TTGTGAGGCCAGGTTGGAGTGGTGGCGAGGTGGGCGCAAGGCTTTCATGGCAAC 3637
Db 117----- 117
QY 3638 GCCACGTGTCCGTCGCCGCCAGGGGTGATCCTGTGCGCTATGCGCCGCGCTGGCGC 3697
Db 118----- 129
QY 3698 GAGCAGAGCGCTTCTCGTGTCCACCTTGGCAACTTGGGCTGGGCAAGTCTCGTG 3757
Db 130 GluGlnArgArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysSerLeu 149
QY 3758 GAGCAGTGGGTGACGAGGCGCGCTCTGCTTGTGCGCTTCCGCCACCAAGCCGCTG 3817
Db 150 GluGlnTrpValThrGluGluAlaAlaCysLeuGlyLysSerLeu----- 166
QY 3818 GGTGATGGCAGAGGGCACAAGCGGAACTGGGAAGCGGGGACGGAAGGCAAC 3877
Db 166----- 166
QY 3878 CTTTACCGGATCTCCCAACCCCGAGGAGCGCCCTTTCGCCCAACCGCCCTTTGGCAA 3937
Db 167----- 180
QY 3938 AGCGTGAGCAACCTGTATCGCTCCCTCAGCTGGGCGCGCTTCCAGTACGACGACC 3997
Db 180 alaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTrpAspAsp 200
QY 3998 TCGCTTCTCAGGCTGTGGACCTAGCTCAGGAGGACTGAAGGAGGAGTCCGGCTTCT 4057
Db 200 oArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluSerGlyPheLe 220
QY 4058 GCGGAGGTGGGACGAGACCGAGAGTCTCTGCAAGGGCGAGTCTCTGAGAGGTGCC 4117
Db 220 uArgGlu----- 222
QY 4118 GGGCTGAGCTGGGGCTCCGAGGGCAGGATTTTCATAGATGGTTTGGAAAGGACAT 4177
Db 222----- 222
QY 4178 TCCAGGAGACCCCACTGTAAAGAGGGCTTGGAGGAGGGGAGACATCTCAGACATGGTGC 4237
Db 222----- 222

QY 4238 TGGAGAGGTGTGCCCCGGGTGAGGGGCGACACGAGAGGCCAAGGACTCTGTACCCCGT 4297
Db 222 ----- 222
QY 4298 CCACGTTGGAGATTTCGATTTTAGTTTCTCCTCTGGCAAGGAGAGAGGGTGGAGGCTG 4357
Db 222 ----- 222
QY 4358 GCATTTGGGAGGGACTTGTGTAGGTGAGTGTAGGTAAAGACAGGAGGCCCTGGGTCTACCT 4417
Db 222 ----- 222
QY 4418 GGAGATGGCTGGGGCCCTGAGACTTGTCCAGGTGAACGCAGAGCACAGGAGGGATTGAGAC 4477
Db 222 ----- 222
QY 4478 CCGGTTCTGTCTGTGTGTAGTGTGAATGCTGCTCCCGTCTCTGCACATCCAGCGCT 4537
Db 223 -----ValLeuAenAlaValProValLeuHisIleProAlaLe 236
QY 4538 GGCTGCAAGGTCTCTACGCTTCCAAAAGGCTTTCCCTGACCCAGCTGGATGAGCTGCTAAC 4597
Db 236 uAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGAGCAGAGTACCTGGGACCCAGCCAGCCAGCCAGCCAGCCAGCTGACTGAGCGCTTCCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GGCAGAGAGGAGAGGTGAGTGTGCTGCCAGCTGGGGGGCAAGGGTGGTGGTTGAA 4717
Db 276 uAlaGluMetGluLys----- 281
QY 4718 CGTCCCAGGAGGAATGAGGGAGGCTGGGCAAAAGGTTGGACCAGTGCATCACCCGGCGA 4777
Db 281 ----- 281
QY 4778 GCCGCATCTGGGCTGACAGGTGCAGAAATTGGAGGTCAATTTGGGGGTACCCCGCTTCTATC 4837
Db 281 ----- 281
QY 4838 CCTGTAGTATCTCTCGGCCCTGCTCAGGCCAAGGGAGGCCCTGAGAGAGAGCTTCAATGA 4897
Db 282 -----AlaLysGlyAsnProGluSerSerPheAsnAs 292
QY 4898 TGAGAACCTGCGCATAGTGTGGTAACTGTTCTTTCGCGGATGCTGACACACCTCGAC 4957
Db 292 pGluAsnLeuArgIleValAlaAspPheSerAlaGlyMetValThrThrSerTh 312
QY 4958 CACGTGGCTGGGGCTCTCTCATGATCTCATACCTGGATGTGACGCGTGAGCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGln----- 328
QY 5018 CTGGGGCCCAAGCAGGAGTGGGAGGAGGTACAGCTGGGGGGCCCTGGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCCGGGCTTCCAGCACAGCGTGGCCAGGCTCTGTAGCGCTTAATTCCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGGAGGAAGGAGAGTGTCCCTGGGTGCTGACCCATTTGGGGAGCGATGTCTGT 5197
Db 328 ----- 328
QY 5198 CAGTCCGTGTCACAGAGATCGACGACGTCGATAGGCGAGTGGCGGACACAGAGATG 5257
Db 329 -Arg--ArgValGlnGlnGluIleAspAspValIleGlyGlnValArgArgProGluMet 347
QY 5258 GGTGACGAGCTCACATGCCCTACACCACTGCCGTGATTCACAGAGTGGAGCGCTTGGG 5317
Db 348 GlyAspGlnAlaHisMetProTyThrThrAlaValIleHisGluValGlnArgPheGly 367

QY 5318 GACATCATCCCCCTGAGTGTGACCCCATATGACATCCCGTGACATCGAAGTACAGGGCTTC 5377
Db 368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CCGATCCCTAAGGTAGGCGCTGGCGCCTCTCACCCAGCTCAGCACCCAGCCTGGTGA 5437
Db 388 ArgIleProLys----- 391
QY 5438 TAGCCCCAGCATGGCTTACTGCCAGGTGGGCCCACTCTAGGAACCTTGGCCACCTAGTCT 5497
Db 391 ----- 391
QY 5498 CAATGCCACACACTGACTGTCCCACTTTGGGTGGGGGTCCAGAGTATAGGCAGGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTCCATCCAGAGCCCCCGTCTAGTGGGGAGACAAACAGGACCTGCCAGAAATGTTG 5617
Db 391 ----- 391
QY 5618 GAGGACCCAGCGCTGCAGGAGAGGGGCGAGTGTGGGTGCTCTGAGAGGTGTGACTGC 5677
Db 391 ----- 391
QY 5678 GCCCTGCTGTGGGTGCGAGAGGGGTACTGTGTGAGCTTCTCGGGCGCAGACTAGTTGACA 5737
Db 391 ----- 391
QY 5738 GAGTCCAGCTGTGTGCCAGGAGTGTGTGTCCCCCGTGTGTTGGTGGCAGGGTCCCAG 5797
Db 391 ----- 391
QY 5798 CATCTTAGAGTCCAGTCCCACTCTCACCTTCATCTCTGCCAGGAAACGACACTCAT 5857
Db 392 -----Gly-ThrThrLeuIle 396
QY 5858 CACCAACCTGTCTCGTCTGAGAGTGAAGCCGTCTGGAGAAAGCCCTTCGCTTCCA 5917
Db 396 eThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHi 416
QY 5918 CCCCAGAACACTTCTCGATGCGCCAGGGCCACTTTGTGAAGCCGAGGCGCTTCTGCTTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPh 436
QY 5978 CTCAGAGTGTCTGTGGGAGCCCGGCTCCCTGTCTCCCTTCCGTGGAGTCTTTCAGGGG 6037
Db 436 e--Ser----- 437
QY 6038 TATCACCCAGGAGCAGGCTCACTGACGCCCTCCCTCCACAGGCCCGCGTGCATGC 6097
Db 438 -----AlaGlyArgArgAlaCys 443
QY 6098 CTGGGGAGCCCTCGCCCGCATGAGCTTCTCTTCTTTCACCTCCCTGCTGCGAGCAC 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeuLeuGlnHis 463
QY 6158 TTCAGCTTCTCGTGGCCCGCCAGCCAGCCCGCCAGCACTCTCGTGTGCTGCTGCTTT 6217
Db 464 PheSerPheSerValProThrGlnProArgProSerHisHisGlyValPheAlaPhe 483
QY 6218 CTGTGACCCCACTCCCTTACGAGCTTTGTGTGTGCCCGCCG 6259
Db 484 LeuValSerProSerProTyGluLeuCysAlaValProArg 497
RESULT 11
ABB09563
ID ABB09563 standard; protein; 497 AA.
XX
AC ABB09563;
XX 02-SEP-2002 (first entry)
DT Human CYP2D6 protein, SEQ ID NO:3.
XX

XX Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;
KW chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;
KW antiarrhythmic; arrhythmia; adrenoceptor antagonist; hypertension;
KW tricyclic antidepressant; procainamide; drug induced lupus syndrome;
KW environmentally linked disease; Parkinson's disease; haplotyping;
KW genotyping; haplotype; genetic variant; single nucleotide polymorphism;
KW SNP; drug screening; drug discovery.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Misc-difference 7
FT /note= "Encoded by RTG in the sequence given in ABQ72215.
FT This residue is Met (encoded by ATG) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT
FT Misc-difference 11
FT /note= "Encoded by RTG in the sequence given in ABQ72215.
FT This residue is Met (encoded by ATG) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT
FT Misc-difference 34
FT /note= "Encoded by YCA in the sequence given in ABQ72215.
FT This residue is Ser (encoded by TCA) rather than Pro
FT (encoded by CCA) in a polymorphic variant"
FT
FT Misc-difference 88
FT /note= "Encoded by CRC in the sequence given in ABQ72215.
FT This residue is His (encoded by CAC) rather than Arg
FT (encoded by CGC) in a polymorphic variant"
FT
FT Misc-difference 91
FT /note= "Encoded by MTG in the sequence given in ABQ72215.
FT This residue is Met (encoded by ATG) rather than Leu
FT (encoded by CTG) in a polymorphic variant"
FT
FT Misc-difference 94
FT /note= "Encoded by CRC in the sequence given in ABQ72215.
FT This residue is Arg (encoded by GCG) rather than His
FT (encoded by CAC) in a polymorphic variant"
FT
FT Misc-difference 98
FT /note= "Encoded by ACS in the sequence given in ABQ72215"
FT
FT Misc-difference 104
FT /note= "Encoded by GVG in the sequence given in ABQ72215.
FT This residue is Ala (encoded by GCG) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT
FT Misc-difference 107
FT /note= "Encoded by WYC in the sequence given in ABQ72215.
FT This residue is Phe (encoded by TTC) rather than Thr
FT (encoded by ACC) in a polymorphic variant"
FT
FT Misc-difference 109
FT /note= "Encoded by RTC in the sequence given in ABQ72215.
FT This residue is Val (encoded by GTC) rather than Ile
FT (encoded by ATC) in a polymorphic variant"
FT
FT Misc-difference 111
FT /note= "Encoded by GGY in the sequence given in ABQ72215"
FT
FT Misc-difference 112
FT /note= "Encoded by TTV in the sequence given in ABQ72215"
FT
FT Misc-difference 120
FT /note= "Encoded by WTC in the sequence given in ABQ72215.
FT This residue is Ile (encoded by ATC) rather than Phe
FT (encoded by TTC) in a polymorphic variant"
FT
FT Misc-difference 128
FT /note= "Encoded by YGG in the sequence given in ABQ72215.
FT This residue is Arg (encoded by GCG) rather than Trp
FT (encoded by TGG) in a polymorphic variant"
FT
FT Misc-difference 136
FT /note= "Encoded by RTS in the sequence given in ABQ72215.
FT This residue is Ile (encoded by ATC) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT
FT Misc-difference 151
FT /note= "Encoded by SAG in the sequence given in ABQ72215.
FT This residue is Glu (encoded by GAG) rather than Gln
FT (encoded by CAG) in a polymorphic variant"
FT
FT Misc-difference 155
FT /note= "Encoded by RAG in the sequence given in ABQ72215.
FT This residue is Lys (encoded by AAG) rather than Glu

FT Misc-difference 232 (encoded by GAG) in a polymorphic variant"
FT
FT Misc-difference 338 /note= "Encoded by CAY in the sequence given in ABQ72215"
FT
FT /note= "Encoded by RTG in the sequence given in ABQ72215.
FT This residue is Met (encoded by ATG) rather than Val
FT (encoded by GTG) in a polymorphic variant"
FT
FT Misc-difference 344
FT /note= "Encoded by YCA in the sequence given in ABQ72215.
FT This is a stop codon (TGA) rather than Arg (encoded by
FT CGA) in a polymorphic variant"
FT
FT Misc-difference 361
FT /note= "Encoded by CAY in the sequence given in ABQ72215"
FT
FT Misc-difference 486
FT /note= "Encoded by ASC in the sequence given in ABQ72215.
FT This residue is Thr (encoded by ACC) rather than Ser
FT (encoded by AGC) in a polymorphic variant"
FT
FT Misc-difference 497
FT /note= "Encoded by YGC in the sequence given in ABQ72215.
FT This residue is Cys (encoded by TGC) rather than Arg
FT (encoded by CGC) in a polymorphic variant"
FT
XX WO200238589-A2.
XX
XX 16-MAY-2002.
XX
XX 09-NOV-2001; 2001WO-US047396.
XX
XX 09-NOV-2000; 2000US-0247943P.
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Anastasio AE, Chew A, Choi JY, Denton RR, Nandabalan K;
XX Petersen N, Rounds E;
XX WPI; 2002-519292/55.
XX N-PSDB; ABQ72215, ABQ72216, ABQ72364.
XX
XX Novel genetic variants of Cytochrome P450, Subfamily IID, Polypeptide 6
XX isogenes, useful for improving efficiency and reliability in drug
XX development for treating hypertension, arrhythmias and Parkinson's
XX disease.
XX
XX Claim 29; Fig 3; 158pp; English.
XX
XX The invention relates to a method for haplotyping the cytochrome P450,
XX subfamily IID, polypeptide 6 (CYP2D6) gene (ABQ72215, ABQ72364) of an
XX individual, and also describes 29 novel polymorphic sites within the
XX human CYP2D6 gene. The CYP2D6 gene is located on chromosome 22q13.1 and
XX contains 9 exons which encode a 497 amino acid protein (ABB09563). CYP2D6
XX is a mono-oxygenase involved in the detoxification of many drugs and
XX environmental chemicals. It plays a role in the metabolism of drugs such
XX as antiarrhythmics, adrenoceptor antagonists and tricyclic
XX antidepressants, and is also involved in the formation of a metabolite
XX linked to the drug-induced lupus syndrome observed with procainamide.
XX Variations in CYP2D6 activity or expression may also influence an
XX individual's susceptibility to environmentally-linked diseases, and it
XX has been demonstrated that CYP2D6 activity may be involved in the
XX pathogenesis of Parkinson's disease, with individuals with a less active
XX form of the enzyme tending to have an earlier onset of this condition.
XX CYP2D6 nucleic acid sequences are useful in studying the expression and
XX function of CYP2D6, and in expressing CYP2D6 protein for use in screening
XX drugs for the treatment of CYP2D6-associated diseases (e.g.,
XX hypertension, atrial and ventricular arrhythmias, Parkinson's disease,
XX and drug-induced lupus syndrome) or which are metabolised by CYP2D6.
XX CYP2D6 nucleic acids and proteins are also useful in studying the effect
XX of polymorphisms on the biological activity of CYP2D6. Polymorphisms in
XX the target region may be determined by the use of allele-specific
XX oligonucleotides (ASOs; ABQ72217-ABQ72303) as probes and primers, and by
XX primer extension using oligonucleotide primers comprising sequences
XX ABQ72304-ABQ72361. The method of the invention is useful for haplotyping
XX the CYP2D6 gene in populations and in individuals, enabling decisions to
XX be made as to whether CYP2D6 is a likely therapeutic target for a disease

CC of interest, and to control for genetically-based bias in the design of
CC drugs that target or are metabolised by CYP2D6. In addition, transgenic
CC animals comprising a human CYP2D6 gene are useful for studying the
CC expression of CYP2D6 isoenzymes in vivo, for in vivo screening and testing
CC of drugs targeted to or metabolised by CYP2D6, and for testing the
CC efficacy of therapeutic agents and compounds for treating CYP2D6-
CC associated conditions in a biological system. The present sequence
CC represents the specifically claimed human CYP2D6 protein. This sequence
CC contains 18 polymorphic sites caused by polymorphisms in the coding
CC sequence (ABQ72216)
xx
SQ Sequence 497 AA;

Alignment Scores:

Prod. No.: 1.99e-106 Length: 497
Score: 1798.00 Matches: 465
Percent Similarity: 34.14% Conservative: 13
Best Local Similarity: 33.21% Mismatches: 19
Query Match: 9.60% Indels: 903
DB: 5 Gaps: 9

US-09-820-788A-3 (1-10278) x ABB09563 (1-497)

QY 2078 ATGGGCTAGAACACTGGTCCCTGGCCATGATAGTGGCCATCTTCTGCTCTGCTG 2137
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
QY 2138 GACCTGATGACCGGCACCAACCTGGCTGGCTGACGCTACCGCCAGCTCCCTGCCACTG 2197
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
21 AspLeuMetHisArgArgGlnAgtProAlaAlaArgTyProProGlyProLeuProLeu 40
QY 2198 CCGGGCTGGGCAACCTTGCTGATGGACTTCCAGAACACACCATCTGCTTCGACCA 2257
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyCysPheAspGI 60
QY 2258 GGTGAGGGAGGAGTCTTGAGGGCGGCAGAGTCTCTGAGATGCCCCACCACCAAA 2317
Db |
60 n----- 60
QY 2318 CATGGTGGTGGGTTAAACCAACAGCTGGATCAGAACGAGCTGAGAGGGGAGCAGG 2377
Db ----- 60
QY 2378 TTTGGGGACGTTCTCTGGGAGGACATTATATATGATGAGTGAAGACTGGATTTTCCAA 2437
Db ----- 60
QY 2438 AGCCCAAGNAGTAGTAGGCAAGGGCTGGAGTGGAGTGGACTTGGCAGTGGGCGATGC 2497
Db ----- 60
QY 2498 AAGCCCATTTGGCAACATATGTTATGAGGTACAAAGTCCCTTCTGTGACACCAAGGA 2557
Db ----- 60
QY 2558 AAGCCCTTGGGAATGGAAGATGAGTTAGTCTCTGAGTGGCGTTTAAATACGAAATCGAGG 2617
Db ----- 60
QY 2618 ATCAAGGGGGTGCAGTGACCCGGTTCAAACCTTTTGCACCTGTGGGTCTCGGCGCTCACT 2677
Db ----- 60
QY 2678 GCTCACCGGATGGACCATCATCTCGGAATGGATGCTAACTGGGGCCCTCTCGGCAATTT 2737
Db ----- 60
QY 2738 TGGTGACTCTTGCAAGGTGATACCTGGGTGACGATCCAACTGAGTTCTCTCATCACAG 2797
Db ----- 60
QY 2798 AAGGTGTGACCCCAACCCCTGCCCCACGATCAGGAGGCTGGGTCTCTCTCTCCACTGC 2857

Db 60 ----- 60
QY 2858 TCACTCTCTGTTAGCCCCGGGGTCTGTCACAGGTTTCAATAGGACTAGGACTCTGTAGTCTG 2917
Db ----- 60
QY 2918 GGTGATCTCTGTTGACAAAGAGGCCCTGACCTCCTCTGTCAGTTGCGGCGCGCTTCG 2977
Db :::: |||||||
61 -----Leu-ArgArgPheG 66
QY 2978 GGGAGGTGTTACGCTGACCTGGCTGGACCGCGGTGGACCGCGGTGCTGCTCAATGGCTGGCGG 3037
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
66 lyAspValPheSerLeuGlnLeuAlaIleThrProValValLeuLeuAsnGlyLeuAla 86
QY 3038 CCGTGGCGAGGCGATGGTGACCCCGCGCGAGGACACGCGCCAGCCCGCCCTGCGCCCA 3097
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
86 laValArgGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProProValProI 106
QY 3098 TCTACCAAGTCTCTGGGCTTGGCGCGCTTCCCAAGCGCGGTGGGGGACAGAGA 3157
Db ||:::|||||||||||||||||||||||||||||||||||||||||
106 leThrGlnIleLeuGlyPheGlyProArgSer-Gln----- 117
QY 3158 CCGCGTTCCTGGTGGGCGCGGTGGACAGTGCAGCTAGCCCAAGCAGCGCCGACGGCG 3217
Db ----- 117
QY 3218 TGGGGTCTCTGGACGTGAACAGAGATAAAGGCCAGCGAGTGGGCTGAGGACAGTGGGCCA 3277
Db ----- 117
QY 3278 GGAAACCACTGCACGGGGAGGTGCGAGTCTGTGGGTGGGAGGGCGGGCTACTGC 3337
Db ----- 117
QY 3338 CCAGACCCCGCAGAGCCCGGTGGCGGAGGCTGATGCGTGAAGTGGCGGTGGCGGGGAC 3397
Db ----- 117
QY 3398 CGCGCTATGCTCGGGCTCAGTGTGGCGGAGCGGGCGGGATCTTCTTGTAGTGAAG 3457
Db ----- 117
QY 3458 GTGTCAGGTGGCGCAGAGACGAGTGGGGCCAAACCCCGCCCGCAGGAGGAGCAATG 3517
Db ----- 117
QY 3518 TGGTGAACAAGAGTGGGCCCTGTGCCAGCTGGACCGGGCTAGGGAATGCGGGAGACC 3577
Db ----- 117
QY 3578 TTGTGAGCGCCAGGTTGGAGTGGGTGGCGAGGTGGGGCCAAAGCCCTTCATGGCAAC 3637
Db ----- 117
QY 3638 GCCCAGTGTCCGTCGCCGCCCGCAGGGGTGATCTGTGCGCTATGGGCGCGCTGGCGC 3697
Db |||||||
118 -----GlyValPheLeuAlaArgTyGlyProAlaT-PArg 129
QY 3698 GAGCAGAGCGCTTCTCGTGTCCACTTGGCCAACTTGGGCTGGGGCAAGAGTCTGCTG 3757
Db |||||||
130 GluGlnArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyIleGlySerLeu 149
QY 3758 GAGCAGTGGTCCAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3817
Db |||||||
150 GluGlnIleValThrGluGluAlaIleCysLeuGlyAlaIlePheAlaAsn----- 166
QY 3818 GGTGATGGCGAGAGGCGCAAAAGCGGAACTGGGAGGGCGGGGAGCGAGGAAGCAACC 3877
Db ----- 166
QY 3878 CTTTACCGCATCTCTCCACCCCGCAGGAGCGCCCTTTTGGCCCAACAGCGCTCTTGGACAA 3937
Db |||||||
167 -----His-Ser-----GlyArgProPheArgProAsnGlyLeuLeuAspI 180

QY 3938 AGCCGTGAGCAACGTGATCGCTCCCTCACCTCGCGCGCGCTTCGAGTACGACGCC 3997
Db 180 sAlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyAspPr 200
QY 3998 TCGCTTCTCCTCAGGCTGCTGAGCTAGCTCAGGAGGACTGAAGAGGAGTTCGGCTTCT 4057
Db 200 oArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuIysGluSerGlyPheLe 220
QY 4058 GCGCAGGTGCGGAGCGAGACCGAGGAGTCTCTGCGGGCGAGCTCCTGAGAGGTGCC 4117
Db 220 uArgGlu 222
QY 4118 GGGGCTGGACTGGGCTCCGAGGCGCAGGATTTGCATAGATGGGTTGGGAAAGGACAT 4177
Db 222 222
QY 4178 TCCAGGAGACCCCACTGTAAAGAGGCGCTGGAGGAGGAGGACATCTCAGACATGGTGC 4237
Db 222 222
QY 4238 TGGGAGAGGTGTCGCCGGGTACGGGGGCACAGSAGAGGCCAAGACTCTGTACCCCGT 4297
Db 222 222
QY 4298 CCACGTTGGAGATTTCGATTTTAGGTTTCTCTCTGGGCAAGGAGAGGGTGGAGGCTG 4357
Db 222 222
QY 4358 GCACTTGGGGGAGCTTGGTGAGGTGAGTGTGTAAGGACAGCAGGCGCTGGGTCTACT 4417
Db 222 222
QY 4418 GGAGATGGCTGGGGCTGAGACTTGTCCAGGTGAACGCGAGACACAGGAGGATTGAGAC 4477
Db 222 222
QY 4478 CCCGTTCTGCTGTTAGGTGCTGAATCTCTCCCGCTCCCTCGACATCCCGCGCT 4537
Db 223 ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe 236
QY 4538 GGCTGGCAAGTCTACGCTTCCAAAAGCTTTCCTGACCCAGCTGGAGTGAAGTCTAAC 4597
Db 236 uAlaGlyIysValLeuArgPheGlnIysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGAGCACAGGATGACTGGGACCCAGCCAGCCACCCGAGACCTGACTGAGGCGCTTCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GGCAAGAGGAGAGGTGAGAGTGGCTGCCAGGTGGGGGCAAGGGTGGTGGTTGAA 4717
Db 276 uAlaGluMetGluIys 281
QY 4718 CGTCCAGGAGGAATGAGGGGAGGCTGGGCAAAAGGTTGGACAGTGATCACCOCGCGA 4777
Db 281 281
QY 4778 GCCGCATCTGGGCTCACAGGTGCAGAAATTGGAGGTCAATTGGGGCTACCCCGTTCTATC 4837
Db 281 281
QY 4838 CCCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCGTTCAATGA 4897
Db 282 AlalysGlyAsnProGluSerPheAsnAs 292
QY 4898 TGAGAACCTGCGCATAGTGGGTAACTGTCTCTGCGGGATGGTGACCACTCGAC 4957
Db 292 pGluAsnLeuArgIleValAlaAspLeuPheSerAlaGlyMetValThrThrSerTh 312
QY 4958 CACGCTGGCTGGGGCTCTCTGCTCATGATCTTACCTGTGATGTCAGCGTGGAGCCGAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGln 328

QY 5018 CTGGGGCCCAAGGCGAGGACTGAGGGAGGAAGGTTACAGCTGGGGGCCCTTGGGCTTAGC 5077
Db 328 328
QY 5078 TGGGACACCCGGGGCTTCCAGCACAGGCGTGGCCAGGCTCCTGTAAAGCTTAACCTTCTCC 5137
Db 328 328
QY 5138 AACACAGAGGAAGAGAGTGTCCCTCGGTGTGACCCATTGTGGGGACGCAATGCTGT 5197
Db 328 328
QY 5198 CCAGTCCGTCTCCACAGGAGATCGACAGCTGATAGGCGAGGTGCGGCGACCGAGATG 5257
Db 329 ArgValGlnGlnGluIleAspAspValIleGlyGlnValArgArgProGluMet 347
QY 5258 GGTGACCAAGGCTCACATGCCCTTACACCACTGCCGTGATTTCACGAGGTGCAGCGCTTTCGG 5317
Db 348 GlyAspGlnAlaHisMetProTyThrThrAlaValIleHisGluValGlnArgPheGly 367
QY 5318 GACATATCCCTCAGGTGTGACCCATATGATCATCCCGTGACATCGAAGTACAGGCTTC 5377
Db 368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CGCATCCCTAAGGTAGGCTGGCGCCCTCTCACCCAGCTCAGCACACCGCTGGTGA 5437
Db 388 ArgIleProIys 391
QY 5438 TAGCCCCAGCATGGCTACTGCGAGGTGGGCCCACTCTAGGAACCCCTGGCCACCTAGTCT 5497
Db 391 391
QY 5498 CAATGCCACCACTGACTGTGCCCACTTGGGTGGGGGTCCAGAGTATAGCGAGGCTG 5557
Db 391 391
QY 5558 GCCTGTCCATCCAGAGCCCCGCTCTAGTGGGGAGACAAACCAGGACCTGCCAGAAATGTTG 5617
Db 391 391
QY 5618 GAGGACCCAGCGCTGCAGGAGAGGGGGCAGTGTGGGTGCTCTGAGAGGTGTGACTGC 5677
Db 391 391
QY 5678 GCCCTGTGTGGGTTCGAGAGGGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACA 5737
Db 391 391
QY 5738 GAGTCCAGCTGTGTGCCAGGCGAGTGTGTCCCGCTGTGTGTGGTGGCAGGGGTCCAG 5797
Db 391 391
QY 5798 CATCTAGATCCAGTCCCACTCTCACCTGTCATCTCTGCCAGGGAACGACACTCAT 5857
Db 392 Gly-ThrThrLeuI 396
QY 5858 CACCAACTGTCTCGTGTGAAGGATGAGGCGCTGTGGAGAGAGCCCTTCCGCTTCCA 5917
Db 396 eThrAsnLeuSerValLeuIysAspGluAlaValTrpGluIysProPheArgPheHi 416
QY 5918 CCCCAGACACTTCTGATGCCAGGGCCACTTTGTGAAGCCGGAGGCTTCTTCCCTTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValIysProGluAlaPheLeuProPh 436
QY 5978 CTCAGCAGGTGCTGTGGGGAGCGCGCTCCTCTGTCCTTCCCTGGAGTCTTTCAGGGG 6037
Db 436 e-Ser 437
QY 6038 TATCACCCAGGAGCGAGGCTCAGCGCCCTCCCTCCACAGGCGCGCTGCATGC 6097
Db 438 438
QY 6098 CTCGGGAGCCCTGGCGCGCATGGAGCTCTTCTTCTTTCACCTCCTCTGTCGAGCAC 6157

Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeuLeuGlnHis 463
6158 TTCAGTCTTCTCCGTGGCCGCGGACAGCCCGCCAGCCACTCTCGTGTGCTCAGCTTT 6217
464 PheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyValPheAlaPhe 483
6218 CTGGTACCCATCCCTCCAGAGCTTTGCTGTGCTGCCCGC 6259
484 LeuValSerProSerProTyrGluLeuCysAlaValProArg 497

RESULT 12
AA026405
ID AAO26405 standard; protein; 497 AA.
XX AC AAO26405;
XX DT -30-JAN-2003 (first entry)
XX DE Human drug-metabolising enzyme related protein.

XX KW Cytostatic; antiparkinsonian; gene therapy; cancer; Parkinson's disease;
KW human drug-metabolising protein; enzyme.
XX OS Homo sapiens.
XX WO200279233-A1.
XX PD 10-OCT-2002.

XX PF 01-APR-2002; 2002WO-US009738.
XX PR 30-MAR-2001; 2001US-00820788.
XX PA (PEKE) PE CORP NY.
XX PA (DFRA/) DI FRANCESCO V.
XX PA (BEAS/) BEASLEY E M.
XX PI Shao W, Yan C;
XX WPI; 2003-040649/03.

XX PT New human drug-metabolizing proteins and nucleic acids related to the
PT Cytochrome P450 IID drug-metabolizing enzyme subfamily, useful for
PT treating a condition mediated by a human enzyme protein e.g., cancer.
XX Disclosure; Fig 2B; 72pp; English.

XX CC The invention relates to a novel isolated polypeptide comprising a 446-
CC amino acid sequence or its allelic variant, orthologue or fragment. The
CC allelic variant or orthologue is encoded by a nucleic acid that
CC hybridises under stringent conditions to the opposite strand of the
CC nucleic acid comprising a sequence having 1537 or 10278 bp. The fragment
CC comprises at least 10 contiguous amino acids of the 446-amino acid
CC sequence. The polypeptide is useful for preparing a pharmaceutical
CC composition for treating a disease or condition mediated by a human
CC enzyme protein, e.g. cancer or Parkinson's disease. This sequence
CC represents a human drug-metabolising related protein of the invention

XX SQ Sequence 497 AA;
Alignment Scores:
Pred. No.: 1,99e-106 Length: 497
Score: 1798.00 Matches: 465
Percent Similarity: 34.14% Conservative: 13
Best Local Similarity: 33.21% Mismatches: 19
Query Match: 9.60% Indels: 903
DB: 6 Gaps: 9
US-09-820-788A-3 (1-10278) x AAO26405 (1-497)

QY 2078 ATGGGGCTAGAGACTGTGTCCTGGCCATGATAGTGGCCATCTTCTGCTCTGTTG 2137
|||||

Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
2138 GACCTGATGCACCGCACCAACGCTGGGTGCACGCTACCCGCCAGGTCCCTGCGCACCTG 2197
|||||
Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaA-gTyrProProGlyProLeuProLeu 40
2198 CCGGGCTGGGCAACCTTGTCTCATGTGGACTTCCAGAACACACATCTGCTTCCACCA 2257
|||||
Db 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAenThrProTyrCysPheAspGI 60
2258 GGTGAGGAGGAGGTCTCTGGAGGGCGGAGAGGTCTCTGAGGATGCCCCACCACGACAA 2317
60 n----- 60
2318 CATGGTGTGTGGTTAAACACACAGGCTGGATCAGAAGCCAGGCTGAGAAGGGGAGCAGG 2377
60 ----- 60
2378 TTTGGGGACGTTCTCTGGGGAGGACATTTATACATGGCATGAAGACTGGATTTTCCAA 2437
60 ----- 60
2438 AGGCCAAGGAAGTAGGGCAAGGCCCTGGAGGTGGAGCTGGCAGTGGCGATGC 2497
60 ----- 60
2498 AAGCCCATTTGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGTGTGACACAGGA 2557
60 ----- 60
2558 AAGGCTTTGGAAATGGAAGATGAGTTAGTCTCTGAGTGGCGTTTAAATCACGAAATCGAG 2617
60 ----- 60
2618 ATGAAGGGGGTGCAGTACCCGGTTCAAAACCTTTTGGACTGTGGGTCTCTGGGGCTCACT 2677
60 ----- 60
2678 GCTCACCAGCATGGACCATCATCTCGGAATGGGATGCTAACTGGGGCCTCTCGGCAATT 2737
60 ----- 60
2738 TGGTGACTCTTCAAGGTATACCTGGGTGACGATCCAAACTGAGTTCTCTCCATCACAG 2797
60 ----- 60
2798 AAGGTGACCCCCACCCCTGCCCCACGATCAGAGGGTGGGTCTCTCTTCCACCTGC 2857
60 ----- 60
2858 TCACCTCTGTAGCCCCGGGGGTCTCCAGGTTCAAATAGGACTAGGACCTGTAGTCTG 2917
60 ----- 60
2918 GGGTATCTGTGCTTGACAAGAGGCCCTGACCCCTCTCGAGTTGGGGCCGCTTCG 2977
61 -----Leu-ArgA-ArgPheG 66
2978 GGGAGCTGTTTCAGCTGCAGCTGGCTGGACGCGGTGTGTGTCTCAATGGGCTGGCGG 3037
66 lyAspValPheSerLeuGlnLeuAlaTrpProValValLeuAenGlyLeuAlaA 86
3038 CCGTGGGAGGCGATGTGTGACCCCGCGGAGGACACGGCCGACCGCCCTGCGCCCA 3097
86 laValArgGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProValProI 106
3098 TCTACCACTCTCGGCTTCGGCGCGCTTCCCAAGGCAAGCGCGGTGGGGGACAGAGA 3157
106 leThrGlnIleuGlyPheGlyProArgSer-Gln----- 117
3158 CCGGCTTTCCGTGGGGCCCCGGGTGGACAGTACCGTAGCCCAAGCAGCGCCGACAGGGCG 3217
117 ----- 117

QY 3218 TGGGGTCTTGGAGTGAACAGAGATAAAGCCACGAGTGGCTGAGGACAGTGGGCCA 3277
Db 117 ----- 117
QY 3278 GGAACACCTGACACGGGGAGGTGAGTCTGTGGGTGGAGGGGGGGCTACTGCG 3337
Db 117 ----- 117
QY 3338 CCAGACCCGCCAGAAAGCCGGTGGCGAGGCTGATGCGTGAAGTGGCGGTGGCGGGGAC 3397
Db 117 ----- 117
QY 3398 CGCGCTATGTCGGGGCTCAGTGTGGCGGACGGCGGGATCTTCTTGTAGTGAAG 3457
Db 117 ----- 117
QY 3458 GTGGTCAGGGTGGCAGAGACGAGGTGGGGCCAAACCCCGCCAGGAGGGGAGCAATG 3517
Db 117 ----- 117
QY 3518 TGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGACCGGGCTAGGGACTGCGGGAGACC 3577
Db 117 ----- 117
QY 3578 TTGTGGAGCCAGGGTTGGAGTGGGTGGCGAGGGTGGGGCCAGGCTTCATGGCAAC 3637
Db 117 ----- 117
QY 3638 GCCCAGTGTCCGTCCCGCCCGAGGGGTATCTGTCGCGTATGGCCCGCGTGGCGC 3697
Db 118 -----GlyValPheLeuAlaArgTyrGlyProAlaTyrArg 129
QY 3698 GAGCAGAGCGCTTCTCCGTGTCCACCTTGGCCAACTTGGGCTGGGCAAGAGTCGCTG 3757
Db 130 GluGlnArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysSerLeu 149
QY 3758 GAGCAGTGGGTACCGAGAGCGCCCTGCTTGTGGCTTCCCGCTCCCGCCAGCAGCGGTG 3817
Db 150 GluGlnTrpValThrGluAlaAlaCysLeuCysAlaAlaPheAlaAsn----- 166
QY 3818 GGTGATGGCAGAGGCGCAAAAGCGGAACTGGGAAGCGGGGACGAGAGGCAACC 3877
Db 166 ----- 166
QY 3878 CCTTACCGCATCTCCCAACCCCGAGAGCGCCCTTTCGCCCAACGSCCTTGGACAA 3937
Db 167 -----His-Ser-----GlyArgProPheArgProAsnGlyLeuLeuAspIy 180
QY 3938 AGCCGTGAGCAACGTGATCGCTCCCTCACCTGCGGGGCGCGCTTCGAGTAGCAGACCC 3997
Db 180 AlaValSerAsnValIleAlaSerLeuThrCysGlyArgPheArgPheGluTyrAspAspPr 200
QY 3998 TCGCTTCTCAGGCTGCTGACCTAGCTCAGAGGAGCTGAAGAGGAGTGGGGCTTCT 4057
Db 200 ArgPheLeuArgLeuLeuAspLeuAlaGlnGlyLeuLysGluLysGluSerGlyPheLe 220
QY 4058 GCGGAGGTGGCGGACGAGACCGAGGAGTCTCTGCGAGGCGGAGCTCTCGAGAGGTGCC 4117
Db 220 uArgGlu----- 222
QY 4118 GGGGCTGGACTGGGGCTCCGAAGGGCAGGATTTCATAGATGGGTTTGGGNAAGGACAT 4177
Db 222 ----- 222
QY 4178 TCCAGGAGACCCACTGTGAAGAGGGCTTGGAGGAGGGGAGACATCTCAGACATGGTCG 4237
Db 222 ----- 222
QY 4238 TGGGAGAGGTGTGCCCGGTACGGGGGCACGAGAGGCCAAGGACTCTGTACCCCGGT 4297
Db 222 ----- 222

QY 4298 CCAGTTGGAGATTTCGATTTTAGGTTTCTCTCTGGGCAAGGAGAGAGGGGTGGAGGCTG 4357
Db 222 ----- 222
QY 4358 GCACTTGGGAGGACTTGGTGAAGTCAAGTCAAGTCAAGGAGGAGGCTTGGGTCTACCT 4417
Db 222 ----- 222
QY 4418 GGAGATGGCTGGGCGCTGAGACTTGTCCAGGTGAACGAGAGCACAGGAGGATTTGAGAC 4477
Db 222 ----- 222
QY 4478 CCCGTTCTGTCTGTGTAGGTGCTGAATGCTGCTCCCGTCTCTCTGACATCCCGAGGCT 4537
Db 223 -----ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe 236
QY 4538 GGCTGGCAAGTCTCAGCTTCCAAAGGCTTCTGACCCAGCTGGATGAGTGCCTTAAC 4597
Db 236 uAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGAGCAGAGATGACCTGGGACCCAGCCAGCCAGCCAGACCTGACTGAGGCTTCCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GGCAAGAGAGAGGTGAGAGTGGCTGCCACGCTGGGGGCAAGGTGTGGGTTGAA 4717
Db 276 uAlaGluMetGluLys----- 281
QY 4718 CGTCCAGGAGAAATGAGGGAGGCTGGGCAAAAGTTGGACCAGTGCATCACCAGCGCA 4777
Db 281 ----- 281
QY 4778 GCCGATCTGGGCTGACAGGTGACAAATTGGAGTCAATTTGGGGCTACCCGTTCTATC 4837
Db 281 ----- 281
QY 4838 CCCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGAGCCCTGAGAGACGCTTCAATGA 4897
Db 282 -----AlaLysGlyAsnProGluSerPheAsnAs 292
QY 4898 TGAGAACCTGCGCATAGTGGTAACTTCTTCCCGGATGTTGACCACTCGAC 4957
Db 292 pGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetValThrThrSerTh 312
QY 4958 CACGCTGGCTGGGGCTCTCTCATGATCTTACACTGATGTCGAGCTGAGCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuMetIleLeuHisProAspValGln----- 328
QY 5018 CTGGGGCCCAAGGAGGAGTGGAGGAGGAGTACAGCTGGGGGCGCCCTGGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCCGGGGCTTCCAGCACAGGCGTGGCCAGGCTCTGTAAAGCTTAACCTCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGAGAGAGAGTGTCCCTCGGTGTGACCCATTGTGGGAGCGCATGTCTGT 5197
Db 328 ----- 328
QY 5198 CCAGTCCGTGTCACACAGGAGATCGACAGCTGATAGGCGAGGTGGCGGACGAGATG 5257
Db 329 -Arg--ArgValGlnGlnGluIleAspAspValIleGlyGlnValAlaArgProGluMet 347
QY 5258 GGTGACAGGCTCATGCTCCCTACACCACTCCCGTGAATTCACGAGGTGCAGCGCTTTGGG 5317
Db 348 GlyAspGlnAlaHisMetProTyrThrThrAlaValIleHisGluValGlnArgPheGly 367
QY 5318 GACATCATCCCTGAGTGTGACCATATGATCTCCGCTGATCATCGAAGTACAGGCTTC 5377
Db 368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CGCATCCCTAAGGTAGGCTCGCGCCCTCTCTCACCCGAGCTCAGCACGACCTGGTGA 5437

```
Db 388 ArgileProlys----- 391
QY 5438 TAGCCCCAGCATGGCTACTGCTCCAGGTGGGCCACTCTAGGAACCTTGGCCACTAGTCTCT 5497
Db 391 ----- 391
QY 5498 CAATGCCACCACACTGACTGTCCCCACTTGGGTGGGGGTCCAGAGATATAGGCAGGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTCCATCCAGAGCCCCCGTCTAGTGGGGAGACAAACAGGACCTCCCAATGTTG 5617
Db 391 ----- 391
QY 5618 GAGGACCCAGCGCTGCAGGAGAGGGGCGAGTGTGGGTGCTCTGAGAGGTGTGACTGC 5677
Db 391 ----- 391
QY 5678 GCCCTGCTGTGGGTCCGAGAGGGTACTCTGGAGCTTCTCGGGCGCAGGACTAGTTGACA 5737
Db 391 ----- 391
QY 5738 GAGTCCAGCTGTGTCCAGGACGTGTGTCTCCCCGTGTGTGGTGGCAGGGTCCCGAG 5797
Db 391 ----- 391
QY 5798 CATCTTAGAGTCCAGTCCCCACTCTCACCCCTGCTCTCGCCAGGGAACGACACTCAT 5857
Db 392 ----- 396
QY 5858 CACCAACCTGTCTCATCGTGTCTGAAGATGAGGCGCTCTGGGAGAGCCCTTCCGTTCCA 5917
Db 396 eThraenLeuSerSerValLeuIysaspGluAlaValTrpGluIysProPheArgPhehi 416
QY 5918 CCCCAGAACCTTCTCTGGATGCCAGGGCCACTTTGTGAAGCCGAGGCTTCTCTGCTTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValIysProGluAlaPheLeuProh 436
QY 5978 CTCAGCAGGTGCTGTGGGGAGCCCGCTCTCTGTCCTCTCTGTCCTGAGTCTTGCAGGG 6037
Db 436 e-Ser----- 437
QY 6038 TATCACCCAGGAGCCAGGCTCACTGACGCCCTCTCCCTCCACAGCGCGCTGCATGC 6097
Db 438 ----- 443
QY 6098 CTCGGGAGCCCTGCGCCGATGAGAGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeuLeuGlnHis 463
QY 6158 TTCAGCTTCTCGTGGCCCGCAGCAGCCCGCCGAGCCACTCTCGTGTCTGCTCAGCTTT 6217
Db 464 PheSerPheSerValProThrGlnGlnProArgProSerHisHisGlyValPheAlaPhe 483
QY 6218 CTGGTGACCCCTATCCCTTACGAGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6259
Db 484 LeuValSerProSerProTyrGluLeuCysAlaValProArg 497
```

RESULT 13

ADB25833
ID ADB25833 standard; protein; 497 AA.

AC ADB25833;

XX 20-NOV-2003 (first entry)

DT Human CYP2D6-related protein #3.

DE human; mutant CYP2D6 gene; drug analysis; drug testing.

XX OS Homo sapiens.

XX

```
PN WO2003050282-A1.
XX 19-JUN-2003.
XX 05-DEC-2002; 2002WO-JP012748.
XX 06-DEC-2001; 2001JP-00372548.
XX (TSUR ) TSUMURA & CO.
XX Taniyama M, Ogawa K, Tsuchiya N, Hibino T;
XX WPI; 2003-505401/47.
XX N-PSDB; ADB25778.
XX Genetic polymorphisms of CYP2D6 gene in human population for analysis of
XX drug effect on individual patients and testing of new drugs.
XX Claim 8; Page 46-50; 75pp; Japanese.
XX The invention comprises mutant forms of the human CYP2D6 gene, containing
XX one or more of the following mutations G125A, C1858T, T2874C and C2875T.
XX The mutant human CYP2D6 genes of the invention are useful for analyzing
XX the effect of drugs on individual patients and testing of new drugs. The
XX present amino acid sequence represents a human protein of the invention.
XX SQ Sequence 497 AA;
XX Alignment Scores:
XX Pred. No.: 1.99e-106 Length: 497
XX Score: 1798.00 Matches: 465
XX Percent Similarity: 34.14% Conservative: 13
XX Best Local Similarity: 33.21% Mismatches: 19
XX Query Match: 9.60% Indels: 903
XX DB: Gaps: 9
XX US-09-820-788A-3 (1-10278) x ADB25833 (1-497)
QY 2078 ATGGGGCTAGAGACACTGGTGGCCCTGGCCATGATAGTGGCCATCTTCTCTCTCTG 2137
Db 1_MetGlyLeuGluAlaLeuValProLeuAlaAlaValLeuAlaValLeuLeuLeuVal 20
QY 2138 GACCTGATGCACCGGCACCAACGCTGGGCTGCACGCTACCCGCCAGGTCCCTGCCACTG 2197
Db 21_AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProGlyProLeuProLeu 40
QY 2198 CCGGGCTGGGCAACCTTCTGCTGCATGTGGACTTCCAGAACACACCATCTGCTTGACCA 2257
Db 41_ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspG 60
QY 2258 GGTGAGGGAGGAGGTCTCTGGAGGGCGGAGAGGTCTCTGAGGATGCCCCACCAGCAA 2317
Db 60 n----- 60
QY 2318 CATGGTGTGGGTTAAACACAGGCTGGATCAGAAGCCAGGCTGAGAGGGGAAGCAGG 2377
Db 60 ----- 60
QY 2378 TTTGGGGACGTTCTCTGGGGAAGGACATTTATACATGGCATGAAGACTGGATTTTCCAA 2437
Db 60 ----- 60
QY 2438 AGGCCAAGGAAGTAGGGAAGGGCTGGAGGTGGAGCTGGACTTGGCAGTGGGCATGC 2497
Db 60 ----- 60
QY 2498 AAGCCCATTTGGCAACATATGTTATGAGTACAAAGTCCCTTCTGCTGACACCAAGGA 2557
Db 60 ----- 60
QY 2558 AAGCCCTTTGGGAATGGAAGATGAGTTAGTCTCTGAGTGGCGTGTAAATCAAGAAATCGAGG 2617
Db 60 ----- 60
```

QY 2618 ATGAAGGGGTGCTGAGTACCGCGTTCAAACCTTTTGCACTGTGGGTCTCTCGGCGCTCACT 2677
Db 60 ----- 60
QY 2678 GCTACCGGCATGGACCATCATCTCGGAATGGATGCTAATCGGGCCCTCTCGGCAATTT 2737
Db 60 ----- 60
QY 2738 TGGTACTCTTGCAGAGTCACTACCTGGGTGACGATCCAAACTGAGTTCCTCCATCAG 2797
Db 60 ----- 60
QY 2798 AAGGTGTGACCCCGCCCTGCCACGATCAGGAGGTGGGTCTCTCTCTCTCTCCACCTGC 2857
Db 60 ----- 60
QY 2858 TCACTCTGTAGTCCCGGGGTCTGTCGAAGTTTCAAATAGGACTAGGACCTGTAGTCTG 2917
Db 60 ----- 60
QY 2918 GGGTATCTGTGCTTGACAAGAGGCCCTGACCCCTCCCTCTGCACTTGGCGCGCTTCG 2977
Db 61 ----- -Leu-ArgArgArgPheG 66
QY 2978 GGGAGCTGTTGAGCTGAGCTGGCTGGAGCGCGGTGCTGCTCAATGGGCTGGCGG 3037
Db 66 IysPvalPheSerLeuGlnLeuAlaTrpProValValValLeuAsnGlyLeuAlaA 86
QY 3038 CCGTCGGGAGCGATGTGACCCCGCGGAGGACACGCGCGGCTGCTGCGCCCA 3097
Db 86 IValArgGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProValProI 106
QY 3098 TCTACAGTCTTGGGCTTGGGCGCGGTTCCAAAGGCAAGCGCGGTGGGGGACAGAGA 3157
Db 106 IeThrGlnIleLeuGlyPheGlyProArgSer-Gln- 117
QY 3158 CCGCGTTTCCTGGGCGCGGTGGAGTGCAGTGCAGTCCCAAGCAGCGCGACAGGGCG 3217
Db 117 ----- 117
QY 3218 TGGGTCTGGAAGTGAACAGAGATAAAGCCAGCGAGTGGGTGAGGACAGTGGGCCA 3277
Db 117 ----- 117
QY 3278 GGAACACCTGACCGGGGAGTCCGAGTCTGTGGGTGGAGGGGGGGGGTACTGCG 3337
Db 117 ----- 117
QY 3338 CCAGACCGCCAGAACCCGTTGGCGAGGCTGATCGTTCGAAGTGGCGGTGGCGGGAC 3397
Db 117 ----- 117
QY 3398 CGCGCTATGCTGCGGGCTCAGTGTGGCGGACGCGGGGATCTTCTTGAGTGAAGAAG 3457
Db 117 ----- 117
QY 3458 GTGGTCAGGGTGGGACAGACGAGGTGGGGCCAAACCCCGCCAGGCGGGGAGCAATG 3517
Db 117 ----- 117
QY 3518 TGGGTGACAAAGATGGGCCCTGTGCCAGCTGACCGGGCTAGGACTGCGGGAGACC 3577
Db 117 ----- 117
QY 3578 TTGTGGAGCGCCAGGGTTTGGAGTGGGTGGCGGAGGTGGGGCCAAAGGCTTCATGSCAAC 3637
Db 117 ----- 117
QY 3638 GCCACGATGTCGTCGCCGCCCGCGGGTATCTGTGCGGTATGGGCGCGCGTGGCGC 3697
Db 118 ----- -GlyValPheLeuAlaArgTyrGlyProAlaTrpArg 129

QY 3698 GAGCAGAGCGCTTCTCGTGTCCACCTTGGCAACTTGGCCCTCGGCAAGAAGTCTGCTG 3757
Db 130 GluGlnArgArgPheSerValSerThrLeuArgAsnLeuGlyLeuLysSerLeu 149
QY 3758 GAGCAGTGGTGACCGAGAGGCGCTTGTGTCGCGCTTGGCGACCAAGCCGCTG 3817
Db 150 GluGlnTrpValThrGluAlaAlaCysLeuCysAlaAlaPheAlaAsn- 166
QY 3818 GGTGATGGCAGAGGGCACAAAGCGGAACTGGGAAGCGGGGACGAGAAAGCAACC 3877
Db 166 ----- 166
QY 3878 CTTTACCGCATCTCCACCCCGAGGACGCCCTTTCGCCCCCAACGCGCTTTGGACAA 3937
Db 167 ----- -His-Ser- 180
QY 3938 AGCGTGAAGCAACGTGATCGCTTCCCTCACCTGGGGCGCGCTTCGAGTACGACGACC 3997
Db 180 sAlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTyrAspAspPr 200
QY 3998 TCGCTTCTCAGGCTGTGGACCTAGCTCAGGAGGGAAGAGGAGTCTGGGCTTCT 4057
Db 200 oArgPheLeuArgLeuAspLeuAlaGlnGluGlyLeuLysGluGluSerGlyPheLe 220
QY 4058 CCGCAGGTGGGAGCGAGACCGAGAGTCTCTGCAGGCGGAGCTCTGAGAGGTGCC 4117
Db 220 uArgGlu- 222
QY 4118 GGGCTGGAAGTGGGGCTTCGAAGGGCAGGATTTGCATAGATGGTTTGGAAAGGACAT 4177
Db 222 ----- 222
QY 4178 TCCAGGACCCCACTGTAAAGGGCTTGAGGAGGGGACATCTCAGCATGTGTCG 4237
Db 222 ----- 222
QY 4238 TGGGAGAGGTGTGCCGGTTCAGGGGCAACGAGGAGGCCAAGGACTCTGTACCCCGT 4297
Db 222 ----- 222
QY 4298 CCAGTTGGAGATTTGATTTTAGTTTCTCTCTGGGCAAGGAGAGGGTGGAGGCTG 4357
Db 222 ----- 222
QY 4358 GCATTTGGGAGGAGCTTGTGAGTCTAGTGTAAAGACAGCAGGCGCTGCTACCT 4417
Db 222 ----- 222
QY 4418 GGAGATGGCTGGGCGCTCAGACTTGTCCAGGTGAACGACAGACAGGAGGATTTGAGAC 4477
Db 222 ----- 222
QY 4478 CCCGTTCTGTGTGTTAGTGTGAATGTGTCTCCGCTCTCTCTGCAATCCCAAGCGCT 4537
Db 223 ----- -ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe 236
QY 4538 GGCTGGCAGGCTCTACCTTCCAAAGGCTTCTTGACCCAGCTGGATGAGTCTCTAAC 4597
Db 236 uAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGAGCAGAGGATGACCTGGGACCCAGCCAGCCACCCAGACCTGACTGAGGCGCTTCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GGCAAAGAGAGAGGTGAGGTGGTGCACCGTGGGGGCAAGGGTGTGGGTGAA 4717
Db 276 uAlaGluMetGluLys- 281
QY 4718 CGTCCAGAGGAATGAGGGAGGCTGGGCAAAAGGTTGGACCATCACCAGCGCA 4777
Db 281 ----- 281
QY 4778 GCCGATCTGGGCTGACAGGTGCAGATTGAGGTTCATTTGGGGGCTACCTCCGCTTCTATC 4837

Db 281 ----- 281
QY 4838 CCCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCAATGA 4897
Db 282 ----- -AlaIysGlyAenProGluSerSerPheAsnAs 292
QY 4898 TGAAACCTGCGCATAGTGGTAACTGTTCTTGGCCGGATGGTACACCACTTCGAC 4957
Db 292 pGluAenLeuArgIleValAlaAlaAspLeuPheSerAlaGlyMetValThrSerTh 312
QY 4958 CAGCTGGCTGGGGCTCTGCTCATGATCTTACACCTGGATGTGACGGTGAGCCAG 5017
Db 312 rThrLeuAlaIrpGlyLeuLeuMetIleLeuHisProAspValGln----- 328
QY 5018 CTGGGGCCCAAGCAGGAGCTGAGGGAGGAAGGGTACAGCTGGGGGCCCTGGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCCGGGGCTTCCAGCAAGCGGTGGCCAGGCTCTGTAAAGCCTAACTTCTCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGGAGGAGAGAGTGTCCCTCGGTGCTGACCCATTGTGGGGAGCGATGCTGT 5197
Db 328 ----- 328
QY 5198 CCAGTCCGTGTCCAACAGGAGATCGACGCTGATAGGGCAGGTGGCGGCACAGAGATG 5257
Db 329 -Arg--ArgValGlnGlnGluLeuAspValIleGlyGlnValArgArgProGluMet 347
QY 5258 GGTGACCAAGCTCACATGCCCTACACACTGCGGTGATTACAGAGTGCAGCGCTTTGGG 5317
Db 348 GlyAspGlnAlaHisMetProTyrThrAlaValIleHisGluValGlnArgPheGly 367
QY 5318 GACATCATCCCTGAGTGTGACCCATATGACATCCCGTGACATCGAAGTACAGGCTTC 5377
Db 368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CGCATCCCTAAGGTAGGCTGGGGCCCTCTCACCCAGCTCAGCACCCAGCCTGGTGA 5437
Db 388 ArgIleProlys----- 391
QY 5438 TAGCCCCAGCATGGCTACTGCTCCAGTGGGGCCCACTTAGGAACTCTGGCCACTAGTCTCT 5497
Db 391 ----- 391
QY 5498 CAATGCCACCACTGACTGTCCCACTTGGTGGGGGGTCCAGATATAGCAGGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTCCATCCAGAGCCCCGCTCTAGTGGGGAGACAAACAGGACCTGCCAGATGTTG 5617
Db 391 ----- 391
QY 5618 GAGGACCCAGCGCTGCAGGGAGAGGGGCGAGTGGGTGCTCTGAGAGGTGTGACTGC 5677
Db 391 ----- 391
QY 5678 GCCCTGCTGGGGTCGGAGAGGGTACTGTGAGCTTCTCGGGCGCAGGACTAGTTGACA 5737
Db 391 ----- 391
QY 5738 GAGTCCAGCTGTGTGCCAGGAGTGTGTCTCCCCGTGTGTTGTGGCGAGGGTCCAG 5797
Db 391 ----- 391
QY 5798 CATCTTAGATCCAGTCCCACTCTCACCGTGCATCTCTGTCGCCAGGAACGACATCAT 5857
Db 392 -----Gly-ThrThrLeuIle 396
QY 5858 CACCAACCTGTATCCGTGCTGAAGGATGAGGCCGTCTGGGAGAACCCCTTCGCTTCCA 5917
|||||

Db 396 eThrAenLeuSerSerValLeuLeuAspGluAlaValTrpGluLysProPheArgPheHi 416
QY 5918 CCCCAGAACACTTCTCGATGTCGCCAGGGCCACTTTGTGAAGCGGAGCCCTTCTCGCTTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPh 436
QY 5978 CTCACAGAGTGTCTGTGGGAGCCGGCTCCCTGTCTCCCTTCCGTGGAGTCTTTCAGGGG 6037
Db 436 e--Ser----- 437
QY 6038 TATCACCCAGGAGCCAGGCTCACTGACGCCCTCCCTCCACAGGCCCGCTGCATGC 6097
Db 438 -----AlaGlyArgArgAlaCys 443
QY 6098 CTCGGGAGCCCTGCGCCGATGGAGCTTCTCTTCTTCCACCTCCCTGCTGCAGCAC 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeuLeuGlnHis 463
QY 6158 TTCAGCTTCTCGTGGCGCCGAGCAGCCCGCCAGCCACTCTCGTGTCTGTCAGCTTT 6217
Db 464 PheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyValPheAlaPhe 483
QY 6218 CTGTGTGACCCCATCCCTACGAGCTTGTGTGTGCTGTCGCCCGC 6259
Db 484 LeuValSerProSerProTyrGluLeuCysAlaValProArg 497
RESULT 14
ABU09593
ID ABU09593 standard; protein; 497 AA.
XX
AC ABU09593;
XX
DT 16-JUL-2003 (first entry)
XX
DE Human cytochrome p450 gene CYP2D6, wild-type protein.
XX
KW Human; cytochrome P450; CYP2D6; chromosome 22; SNP; drug metabolism;
cardiovascular disorder; psychiatric disorder; drug sensitivity.
XX
OS Homo sapiens.
XX
PN BP1281755-A2.
XX
PD 05-FEB-2003.
XX
PF 16-JUL-2002; 2002EP-00254972.
XX
PR 31-JUL-2001; 2001US-0309111P.
XX
PA (PFIZ) PFIZER PROD INC.
XX
PI Milos PM, Webb SM;
XX
DR WPI; 2003-373769/36.
DR N-PSDB; ACA61303.
XX
PT New cytochrome P450 2D6 gene variants and polypeptides, useful for
determining if a subject has or is at risk of developing a drug
sensitivity condition or disorder that is associated with an aberrant
CYP2D6 activity.
XX
PS Claim 11; Fig 4; 88pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising a cytochrome
P450 2D6 gene variant, e.g. G5799C or C5916AT (referring to the genomic
sequence or the same variant nucleotide in the corresponding cDNA
sequences). Also included are probes, primers (allele specific
oligonucleotides) and arrays used to detect and/or amplify the CYP2D6
gene polymorphic regions, the variant polypeptides, antibodies which are
capable of distinguishing between the variant and wild-type polypeptides,
determining whether a subject has a genetic deficiency for metabolising a
drug, evaluating therapy with a drug metabolised by P450 CYP2D6 and
determining whether an individual is susceptible to being a poor

CC metaboliser of drugs. The DNA probe is useful for hybridising to a
CC variant form of the CYP2D6 gene. The primer is useful for amplifying the
CC C5816TA allelic variant. The allele specific nucleotide is useful for the
CC detection of the C5816TA allelic variant. The methods are useful for
CC determining whether a subject has a genetic deficiency for metabolising a
CC drug, evaluating therapy with a drug metabolised by P450 CYP2D6, and
CC determining if an individual is susceptible to being a poor metaboliser
CC of drugs. The nucleic acids are useful as probes or primers for
CC determining whether a subject has a genetic deficiency for metabolising
CC drugs that are substrates of P450 CYP2D6. The methods are useful for
CC determining if a subject has or is at risk of developing a drug
CC sensitivity condition or disorder that is associated with an aberrant
CC CYP2D6 activity, e.g. an aberrant level of a CYP2D6 protein or an
CC aberrant CYP2D6 bioactivity. The methods are also useful in selecting the
CC appropriate drugs or determining the course of treatment to administer to
CC a subject to treat cardiovascular or psychiatric disorders, or for
CC treating a subject with a drug sensitivity or disorder associated with a
CC specific allelic variant of a polymorphic region of the CYP2D6 gene. The
CC antibodies are useful for monitoring CYP2D6 protein levels in an
CC individual for determining whether a subject has a disease or conditions
CC associated with an aberrant CYP2D6 protein level. The gene is located on
CC human chromosome 22. The present sequence is the wild-type CYP2D6 protein
XX
SQ Sequence 497 AA;

Alignment Scores:
Pred. No.: 3, 09e-106 Length: 497
Score: 1795.00 Matches: 464
Percent Similarity: 34.14% Conservative: 14
Best Local Similarity: 33.14% Mismatches: 19
Query Match: 9.58% Indels: 903
DB: Gaps: 9

US-09-820-788A-3 (1-10278) x ABU09593 (1-497)

```
QY 2078 ATGGGGCTAGAGCACTGGTGGCCCTGGCCATGATAGTGGCCATCTCTGCTCTCTGGTG 2137
DB 1 MetGlyLeuGluAlaLeuValProLeuAlaValIleValAlaIlePheLeuLeuVal 20
QY 2138 GACCTGATGCACCGGCACCAAGCTGGGGCTGCACGCTACCGCCAGGTCCCTGCGCACTG 2197
DB 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProProGlyProLeuProLeu 40
QY 2198 CCCGGCTGGGCAACCTTGTGCTGCTGATGCTGCTCCAGAACACACCATCTCTTCGACCA 2257
DB 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspG1 60
QY 2258 GGTGAGGGAGGAGGTCTCTGGAGGGCGGCGAGAGTCTCTGAGGATGCCCCACCACGACAAA 2317
DB 60 n----- 60
QY 2318 CATGGGTGGTGGTTAAACACAGAGCTGGATCAGAAAGCCAGGCTGAGAAAGGAGAGCAGG 2377
DB 60----- 60
QY 2378 TTGGGGGACGTTCTCTGGGAAGGACATTTATACATGCATGAGGACTGGATTTTCCAA 2437
DB 60----- 60
QY 2438 AGGCCAAGGAAGTAGTAGGCAAGGGCCTGGAGGTGGAGCTGGACTTGGCAGTGGGCGATGC 2497
DB 60----- 60
QY 2498 AAGCCATTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACACCAAGGA 2557
DB 60----- 60
QY 2558 AAGGCTTTGGGAATGGAAGATGAGTTACTCTGAGTGGCGTTTAAATCAGGAATCGAGG 2617
DB 60----- 60
QY 2618 ATGAAGGGGTGCAGTGAACCGGTTCAAACCTTTTGCACCTGTGGGTCTCTCGGGCTCACT 2677
```

```
DB 60----- 60
QY 2678 GCTCACCGGATGGACCATCATCTCGGGAATGGGATGCTAACTGGGGCCTCTCGGCAATTT 2737
DB 60----- 60
QY 2738 TGGTGACTCTTTCAAGAGTTCATACCTGGGTGAGCGATCCAAACTGAGTTCTTCATCAGAG 2797
DB 60----- 60
QY 2798 AAGGTGTGACCCCAACCCCTGCCCCCAGCATCAGGAGGTGGGTCTCTCTCTTCCACCTGC 2857
DB 60----- 60
QY 2858 TCACCTCTGTGTAGCCCGGGGGTGTCTCAAGGTTCAAATAGGACTAGGACCTGTAGTCTG 2917
DB 60----- 60
QY 2918 GGGTGATCTGTGTTGACAAGAGGCCCTGACCTCCCTCTGCAGTTTGGCGCGCGCTTCG 2977
DB 61-----Leu-ArgArgArgPheG 66
QY 2978 GGGACGTGTTTTCAGCTGCGCTGAGCGCGGTGGTGTCTCAATGGCTGGCGG 3037
DB 66 LyAspValPheSerLeuGlnLeuAlaTrpThrProValValLeuAsnGlyLeuAla 86
QY 3038 CCGTCCGGGAGCGATGGTGAACCGCGCGGAGGACACGCGCCGACCGCGCTGGCGCCA 3097
DB 86 laValArgGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProValProI 106
QY 3098 TCTACCAAGTCTCTGGGCTTCCGGCGCGTTCCTCCAAAGGCAAGCGCGGTGGGGGACAGAGA 3157
DB 106 leThrGlnIleLeuGlyPheGlyProArgSer-Gln----- 117
QY 3158 CCGGCTTTTCTGTGGGCCCCCGGGTGGACAGTGCACCGTAGCCCAAGCAGCGCCGACAGGCG 3217
DB 117----- 117
QY 3218 TGGGGTCTCTGACGTGAAACAGAGATAAAGCCAGCGAGTGGGTGAGGACAGTGGGCGCA 3277
DB 117----- 117
QY 3278 GGAACACACCTGCACGGGGAGGTGCGAGTCTGTGGGTGGGAGGGGGCGGCTACTGC 3337
DB 117----- 117
QY 3338 CCAGACCGCCAGAAAGCCCGGTGGCGGAGGTGATGCTGCGAAGTGGCGGTGGCGGGGAC 3397
DB 117----- 117
QY 3398 CGCGCTTATGCTCGGGGCTCAGTGTGGGCGGGAACGGCGGGATCTTCTTGAGTGGAAAG 3457
DB 117----- 117
QY 3458 GTGGTCAGGTGGGCGAGAGACAGGTGGGGCCAAACCCCGCCCCCAGGCGAGGAGCAATG 3517
DB 117----- 117
QY 3518 TGGGTGAGCAAGAGTGGGGCCCTGTGCCAGCTGGAACGGGCTAGGAGACTCGGGGAGACC 3577
DB 117----- 117
QY 3578 TTGTGGAGCGCCAGGGTTGGAGTGGGTGGCGGAGGTGGGGCCAAAGGCTTTCATGGGCAAC 3637
DB 117----- 117
QY 3638 GCCCAGCTGTCCGTCGCCGCCCGGAGGGGTGATCTGTGCGCTATGSCCGCGCTGGCGC 3697
DB 118-----GlyValPheLeuAlaArgTyrGlyProAlaTrpArg 129
QY 3698 GAGCAGAGCGCTCTCTCGCTGTCCACCTTGGCAACTTGGGCGCTGGGCAAGAGTCTGCTG 3757
DB 130 GluGlnArgArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysLysSerLeu 149
```

Qy	3758	GAGCAGTCGGGTGACCGAGGAGGCGCGCTGCTTGTGTCGCGCTTCGCCGACCAAGCCGGTG	3817
Db	150	GluGlnTrpValThrGluAlaAlaCysLeuCysAlaAlaPheAlaAsn-----	166
Qy	3818	GGTGATGGCAGAGGGCACAAAGCGGAACTGGGAAGCGGGGACGGAGAAGGCAACC	3877
Db	166	-----	166
Qy	3878	CTTTACCCGCACTCTCCCAACCCCAAGACGCGCCCTTCGCGCCCAACGCGCTCTTGACAA	3937
Db	167	-----His-Ser-----GlyArgProPheArgProAsnGlyLeuLeuAspLy	180
Qy	3938	AGCCGTGAGCAACGTGATCGCCTCCCTCACTGCGCGCGCGCGCTTCGAGTACACGACCC	3997
Db	180	sAlaValSerAsnValIleAlaSerLeuThrCysGlyArgArgPheGluTrpAspAspPr	200
Qy	3998	TCGCTTCTCTCAGGCTGCTGGACCTAGCTCAGGAGGGACTGAAGGAGGAGTCGGCTTTCT	4057
Db	200	oArPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuysGluGluSerGlyPheLe	220
Qy	4058	GCAGGAGTGCAGGACGAGACCGAGGAGTCTCTGCAGGCGGAGCTCTCGAGAGGTGCC	4117
Db	220	uArgGlu-----	222
Qy	4118	GGGCGTGGACTGGCGCCTCCGAAGGCGAGNATTGCAATAGATGGGTTTGGAAAGGACAT	4177
Db	222	-----	222
Qy	4178	TCCAGGAGACCCCACTGTAGAAGGGCTGGAGGAGGGGACATCTCAGACATGCTCG	4237
Db	222	-----	222
Qy	4238	TGGAGAGGTGTGCCCGGTACGGGGCAACAGGAGAGGCCAAGGACTCTGTATACCCCGT	4297
Db	222	-----	222
Qy	4298	CCAGTTGGAGATTTCGATTTTAGGTTTCTCTCTGGCAAGGAGAGAGGGTGGAGGCTG	4357
Db	222	-----	222
Qy	4358	GCACTGGGGAGGACTTGGTGAAGTCAAGTAAAGCAGCAGGAGGCCCTGGGCTACCT	4417
Db	222	-----	222
Qy	4418	GGAGATGGCTGGGCGCTGAGACTTGTCCAGGTGAACCGACAGCACAGAGGGGATGAGAC	4477
Db	222	-----	222
Qy	4478	CCCCTTCTGTGTAGTGTGCTGAATGCTGTCTCCCGTCTCTCTGCACATCCAGGCT	4537
Db	223	-----ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe	236
Qy	4538	GGCTGGCAAGTCTACGCTTCCAAAAGGCTTCTGCACCGAGCTGGATGAGCTGTAAAC	4597
Db	236	uAlaGlyIysValLeuArgPheGlnIysAlaPheLeuThrGlnLeuAspGluLeuLeuTh	256
Qy	4598	TGACACAGGATGACTCTGGGACCCAGCCACCCGAGACTGACTGAGCGCTTCCT	4657
Db	256	rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe	276
Qy	4658	GGCAAGNAGGAGAAGGTGAGATGGCTGCCACGTTGGGGGCAAGGGTGGTGGGTTGAA	4717
Db	276	uAlaGluMetGluIys-----	281
Qy	4718	CGTCCAGAGGAATGAGGGGAGGCTGGGCCAAAAGTTGGACCAAGTCATCATCCCGCGCA	4777
Db	281	-----	281
Qy	4778	GCCGGATCTGGGCTGACAGGTGCAGAAATTGGAGGTCATTTGGGGGCTACCCCGTTCTATC	4837
Db	281	-----	281

Qy	4838	CCCTGAGTATCTCTCGGCCCTCTGCTCAGGCCAAGGGGAGCCCTGAGAGCAGCTTCAATGA	4899
Db	282	-----AlaLysGlyAsnProGluSerPheAsnAs	292
Qy	4898	TGAAACCTCGCATAGTGGTGAACCTGTTCTTTCGCCGGATGGTGACCACTCGAC	4957
Db	292	-----AlaLysGlyAsnProGluSerPheAsnAs	312
Qy	4958	CACGCTGGCTCGGGCTCTCTGCTCATGATCTTACACCTGGATGTGTCAGCGTGAGCCACG	5017
Db	312	rThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGln	328
Qy	5018	CTGGGGCCCAAGCAGGGACTGAGGAGGAGGAGGTACAGCTGGGGGCCCTGGGCTTAGC	5077
Db	328	-----	328
Qy	5078	TGGGACACCCGGGCTTCCAGCACAGCGGTGGCCAGGCTCTCTTAAGCCTAACTTCTCTCC	5137
Db	328	-----	328
Qy	5138	AACACAGGAGGAGGAGTGTCCCTGGGTGTGACCCATTGTGGGGACGCATGTCTGT	5197
Db	328	-----	328
Qy	5198	CCAGTCCGTGTCAAACAGGAGATCGACGAGTGAAGGCGAGTGGCGGCACAGAGATG	5257
Db	329	--Arg--ArgValGlnGlnIleAspAspValIleGlyGlnValArgArgProGluMet	347
Qy	5258	GGTGACGAGCTCACATGCCCTACACACTGCCGTGATTCACAGGTGCAGCCCTTTGGG	5317
Db	348	GlyAspGlnAlaHisMetProTyrThrAlaValIleHisGluValGlnArgPheGly	367
Qy	5318	GACATCATCCCTCGTGTGACCCATATGACATCCCGTGACATCGAAGTACAGGGGCTTC	5377
Db	368	AspIleValProLeuGlyMetThrHisMetThrSerArgAspIleGluValGlnGlyPhe	387
Qy	5378	CGCATCCCTAAGTAGGCGCTGGCGCCCTCTCACCACGCTCAGCACACGACCTGTGTGA	5437
Db	388	ArgIleProLys	391
Qy	5438	TAGCCCCAGCATGGCTACTGCCAGGTGGGCCCACTCTAGGAACCTCGGCCACCTAGTCCT	5497
Db	391	-----	391
Qy	5498	CAATGCCACCACACTGACTGTCTCCCACTTGGGTGGGGGTCCAGAGTATAGGACGGGCTG	5557
Db	391	-----	391
Qy	5558	GCCTGTCCATCCAGAGCCCCGTCTAGTGGGGAGACAAACACGACCTGCCAGATGTTG	5617
Db	391	-----	391
Qy	5618	GAGGACCCAGCGCTGCAGGGAGAGGGGCGAGTGTGGGTGCTCTGAGAGGTGTGATGC	5677
Db	391	-----	391
Qy	5678	GCCCTGTGTGGGTGGAGAGGGTACTGTGAGGCTTCTCGGGCGCAGGACTAGTTGACA	5737
Db	391	-----	391
Qy	5738	GAGTCCAGCTGTGTGCCAGGCAGTGTGTGTCCTCCCGTGTGTTTGGTGGCAGGGGTCCCAG	5797
Db	391	-----	391
Qy	5798	CATCTTAGTGCAGTCCCCCACTCTCACCTGCATCTCTGCCACGGGAACGACATCAT	5857
Db	392	-----Gly--ThrThrLeuIle	396
Qy	5858	CACCAACTGTCTCGTGTGAAGGATGAGGCCGTCTGGGAGAAGCCCTTCGCTTCCA	5917
Db	396	eThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHi	416
Qy	5918	CCCCGAACACTCTCTGGATGCCCAGGGCCACTTTGTGAAGCCGGAGGCGCTTCTCGCTTT	5977

Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheVallysProGluAlaPheLeuProPh 436
QY 5978 CTCACAGGTGCTGTGGGAGCCCGGCTCCCTGTCTCCCTCCGTGGAGTCTTTCAGGGG 6037
Db 436 e--Ser----- 437
QY 6038 TATACCCAGGAGCAGGCTCACTGACGCGCCCTCCCTCCACAGCGCGCGTGCATGC 6097
Db 438 -----AlaGlyArgAlaCys 443
QY 6098 CTCGGGAGCCCTGCGCGCATGAGCTTCTCTCTTTCACCTCCTGCTGCAGCAC 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeuLeuGlnHis 463
QY 6158 TTCACCTTCTCGTGGCGCGGAGCAGCGCGCGCCAGCCTCTCGTGTGCTCACTTT 6217
Db 464 PheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyValPheAlaPhe 483
QY 6218 CTGTGACCCATCCCTACGAGCTTGTGCTGTGCGCCCGC 6259
Db 484 LeuValSerProSerProTyrGluLeuCysAlaValProArg 497

RESULT 15

ABR82026

ID ABR82026 standard; protein; 497 AA.

XX ABR82026;

AC ABR82026;

XX 22-SEP-2003 (first entry)

DT Human cytochrome P450 2D6 amino acid sequence.

XX Human; protein array; protein moiety; phenotype; drug discovery;

XX naturally occurring variant; pharmacogenomic; diagnostic;

KW parallel analysis; tumour suppressor; p53; cytochrome P450.

XX Homo sapiens.

OS WO2003048768-A2.

XX 12-JUN-2003.

XX 05-DEC-2002; 2002WO-GB005499.

XX 05-DEC-2001; 2001US-0335806P.

PR 16-SEP-2002; 2002US-0410815P.

XX (SENS-) SENSE PROTEOMIC LTD.

PA Boutell JM, Godber BLJ, Hart DJ, Blackburn JD;

PI WPI; 2003-569063/53.

XX N-PSDB; ACF06056.

XX New protein array, useful for determining the phenotype of a naturally

PT occurring variant of a DNA sequence of interest, comprises a surface upon

PT which at least two protein moieties are deposited.

XX Example 5; Fig 13B; 84pp; English.

XX The present invention describes a protein array comprising a surface upon

CC which at least two protein moieties are deposited at spatially defined

CC locations, where the protein moieties are naturally occurring variants of

CC a DNA sequence of interest. Also described: (1) making a protein array;

CC (2) screening a set of protein moieties for molecules that interact with

CC one or more proteins; and (3) simultaneously determining the relative

CC properties of members of a set of protein moieties. The protein array can

CC be used for determining the phenotype of a naturally occurring variant of

CC a DNA sequence of interest. The protein array is useful for drug

CC discovery, pharmacogenomics and diagnostics. The protein array allows the

CC parallel analysis of closely related proteins with a sensitivity that is

CC at least comparable to existing methods, if not better, with small

CC volumes of potentially expensive ligands, and in a quantitative,
CC comparative functional analysis manner not previously possible. ACF06000
CC to ACF06056 and ABR81975 to ABR82026 represent sequences used in the
CC exemplification of the present invention

XX SQ Sequence 497 AA;

Alignment Scores: 3.09e-106 Length: 497
Pred. No.: 1795.00 Matches: 464
Score: 34.14% Conservative: 14
Percent Similarity: 33.14% Mismatches: 19
Best Local Similarity: 9.58% Indels: 903
Query Match: 6 Gaps: 9
DB: 9

US-09-820-788A-3 (1-10278) x ABR82026 (1-497)

QY 2078 ATGGGGCTAGAACACACTGGTGGCCCATAGTAGGGCCATCTTCTGTCTCTGGTG 2137
Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValleValAlaPheLeuLeuVal 20
QY 2138 GACCTGATGCACCGCACCAACCGCTGGCTGCAGCTTCCAGNACACACATCTCTGCACCA 2197
Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProGlyProLeuProLeu 40
QY 2198 CCGGGCTGGCAACCTTGTCTGCATGTGGACTTCCAGNACACACATCTCTGCACCA 2257
Db 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspG 60
QY 2258 GGTGAGGAGGAGGTCTCTGGAGGGCGGAGAGGTCTGAGGATGCCCCACCACGACAA 2317
Db 60 n----- 60
QY 2318 CATGGTGTGGTGTAAACACAGCGTGGATCAGAACCGAGCTGAGAAAGGAGCAGG 2377
Db 60 ----- 60
QY 2378 TTTGGGGACGTTCTCTGGGGAAGGACATTTATATCATGTCATGAAGGACTGGATTTCCAA 2437
Db 60 ----- 60
QY 2438 AGGCCAAGAAAGAGTAGGGCAAGGCGCTGGAGGTGGAGCTGGACTTGGCAGTGGGCGATGC 2497
Db 60 ----- 60
QY 2498 AAGCCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACACCAAGGA 2557
Db 60 ----- 60
QY 2558 AAGSCCTTGGGAATGGGAAGATGAGTTAGTCTGAGTGGCGTTAAATCAGGAATCGAGG 2617
Db 60 ----- 60
QY 2618 ATGAAGGGGGTGCAGTGCACCGGTTCAACCTTTTGCATGTGGGTCTCTGGGCGCTCACT 2677
Db 60 ----- 60
QY 2678 GCTCACCAGGATGGACCATCATCTGGGAATGGGATGCTAACTGGGGCGCTCTCGGCAATTT 2737
Db 60 ----- 60
QY 2738 TGGTGACTCTTGAAGGTTCATACCTGGGTGAGGATCCAAACTGAGTTCCTTCATCAGAG 2797
Db 60 ----- 60
QY 2798 AAGGTGTGACCCCGCCCGCTGCCCGCAGCATCAGGAGGTGGGTCTCTCTCTCTCCACCTGC 2857
Db 60 ----- 60
QY 2858 TCACTCTCTGTAGCCCGGGGGTCTGCCAAGGTTCAAATAGGACTAGGACCTGTAGTCTG 2917
Db 60 ----- 60

QY 2918 GGGTGATCTGGCTTGACAGAGCCCTGACCTCCCTCTGCGATTGGCGCGCGCTTCG 2977
Db 61 - - - - - Leu-ArgArgPheG 66
QY 2978 GGGACGGTTTCAGCTCAGCTGGACGCGGGTGGTCTCAATGGGCTGGCGG 3037
Db 66 lyaspValPheSerLeuGlnLeuAlaTrpThrProValValLeuAsnGlyLeuAla 86
QY 3038 CCGTGGCGAGGCGATGGTACCCGCGCGAGGACACGGCCGCGCTTCGCGCCA 3097
Db 86 laValArgGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProProValProI 106
QY 3098 TCTACCAAGTCTCGGGCTTCGGCGCGCTTCCCAAGGCAAGCGCGGTGGGGACAGAGA 3157
Db 106 leThrGlnIleLeuGlyPheGlyProArgSer-Gln- - - - - 117
QY 3158 CCGGTTTCGTGGCGCCCGGGTGGACAGTGACGTAGCCCAAGCAGCGCCGACAGGGCG 3217
Db 117 - - - - - 117
QY 3218 TGGGGTCTGGACGTGAACAGAGATAAAGGCCAGCGAGTGGGCTGAGGACAGTGGGCCA 3277
Db 117 - - - - - 117
QY 3278 GGAACCACTGCACGGGGAGGTGCGAGTCTGTGGGCTGGAGAGGGCGGGCTACTGC 3337
Db 117 - - - - - 117
QY 3338 CCAGACCCGCCAGAAGCCCGTGGCGAGGCTGATGGTGAAGTGGCGGTGGCGGGAC 3397
Db 117 - - - - - 117
QY 3398 CGCGCTATGCTGGGGCTCAGTGTGGCGGGACGGCGGGATCTTCTTGAAGTGAAG 3457
Db 117 - - - - - 117
QY 3458 GTGGTCAGGTGGGCAGACAGAGTGGGGCCAAACCCCGCCAGGAGGGAGCAATG 3517
Db 117 - - - - - 117
QY 3518 TGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGGCTAGGACTGCGGGAGACC 3577
Db 117 - - - - - 117
QY 3578 TTGTGAGCGCCAGGGTTGAGTGGGTGGCGAGGGTGGGGCCCAAGGCTTCATGGAAC 3637
Db 117 - - - - - 117
QY 3638 GCCCAGTGTCTCGCCGCCCCAGGGGTGATCTGTGCGCTATGGCCCGCGTGGCGC 3697
Db 118 - - - - - GlyValPheLeuAlaArgTyrglyProAlaTrpArg 129
QY 3698 GAGCAGAGGGCTTCTCCGTGTCCACTTGGCGCAACTTGGGCTGGGCAAGAGTGGCTG 3757
Db 130 GluGlnArgArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysSerLeu 149
QY 3758 GAGCAGTGGGTGACGAGAGCGCGCTGCTTGTGCGCTTGGCGCGCAAGCGCGTG 3817
Db 150 GluGlnTrpValThrGluGluAlaLysLeuCysAlaAlaPheAlaAsn- - - - - 166
QY 3818 GGTGATGGGCAGAGGGCCAAAGCGGAACTGGGAAGGGGGGACGAGAGGCAACC 3877
Db 166 - - - - - 166
QY 3878 CCTTACCGGATCTCCCAACCCAGGAGCGCCCTTTGCGCCCAAGCGGCTCTTGACAA 3937
Db 167 - - - - - His-Ser- - - - - GlyArgProPheArgProAsnGlyLeuLeuAspI 180
QY 3938 AGCCGTGAGCAAGCTGATCGCTCCCTCAGCTCGGGCGCGCTTCAGTACGACGCC 3997
Db 180 saIaValSerAsnValIleAlaSerLeuThrCysGlyArgPheGluTrpAspAspPr 200
QY 3998 TCGCTTCTCAGGCTGTGACCTAGCTCAGGAGGAGTGAAGAGGAGTTCGGCTTCT 4057

Db 200 oArgPheLeuArgLeuLeuAspLeuAlaGlnGlyLeuLysGluGluSerGlyPheLe 220
QY 4058 GCGCGAGGTGCGGAGGAGAGACCGAGGAGTCTCTCAGGGCGAGCTCTCGAGAGGTGCC 4117
Db 220 uArgGlu- - - - - 222
QY 4118 GGGGCTGGACTGGGGCTCCGAAGGCGAGGATTTTCATAGATGGGTTTGGGAAAGGACAT 4177
Db 222 - - - - - 222
QY 4178 TCCAGGAGACCCCACTGTGAAGAGGCGCTGGAGGAGGGGACATCTCAGACATGGTCG 4237
Db 222 - - - - - 222
QY 4238 TGGGAGAGTGTGCCCGGTTCAGGGGGCACAGGAGGCCAAGGACTCTGTACCCCGT 4297
Db 222 - - - - - 222
QY 4298 CCACGTTGAGATTTTCGATTTTTCCTCTCTGGGCAAGGAGAGAGGTGGAGGCTG 4357
Db 222 - - - - - 222
QY 4358 GCACTTGGGAGGAGCTTGGTGAGGTCACTGGTGAAGGACAGGAGGCCCTGGGTCTACCT 4417
Db 222 - - - - - 222
QY 4418 GGAGATGGCTGGGGCTGAGACTTGTCCAGGTGAAACGACAGACAGAGGGGATTGAGAC 4477
Db 222 - - - - - 222
QY 4478 CCGCTTCTGTCTGGTGTAGGTGCTGAATGCTGTCCCGCTCTCTCCACATCCAGCGCT 4537
Db 223 - - - - - ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe 236
QY 4538 GGTGGCAAGGTCTTACGCTTCCAAAGGCTTTCCTGACCCAGCTGGATGAGCTGTAAC 4597
Db 236 uAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TCAGCACAGGATGACTGGGACCCAGCCAGCCAGCCAGCCAGACCTGACTGAGCGCTTCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GCGAAGAGAGAAAGGTGAGAGTGGCTGCCACGGTGGGGGCAAGGGTGGTGGGTGA 4717
Db 276 uAlaGluMetGluLys- - - - - 281
QY 4718 CGTCCAGGAGGAATGAGGGGAGGTGGGCAAAAGTTGGAACAGTGCATCACCCGCGCA 4777
Db 281 - - - - - 281
QY 4778 GCGCATCTGGGCTGACAGGTGCAGAAATTGGAGGTCAATTGGGGGTACCCGTTCTATC 4837
Db 281 - - - - - 281
QY 4838 CCTGTGATCTCTCTCGGCCCTGCTCAGCCCAAGGGAGCCCTGAGAGCAGCTTCAATGA 4897
Db 282 - - - - - AlalysGlyAsnProGluSerSerPheAsnAs 292
QY 4898 TGAGAACCTTGCATAGTGGTGGTAACCTGTTCCTTTCGCGGATGGTGACCACTTCGAC 4957
Db 292 pGluAsnLeuArgIleValAlaAspLeuPheSerAlaGlyMetValThrThrSerTh 312
QY 4958 CACGCTGGCTGGGCTCTCTCTCATGATCTCTACACCTGGATGTGACGCGTGAGCCCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGln- - - - - 328
QY 5018 CTGGGGCCCAAGGAGGAGTGGAGGAGGAGGTACAGCTGGGGGCCCTTGGGCTTAGC 5077
Db 328 - - - - - 328
QY 5078 TGGGACACCCCGGGGCTTCCAGCACAGGGGTGGCCAGGCTCTGTAAAGCCTAACTTCTCTCC 5137

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 25, 2004, 02:15:53 ; Search time 474.102 Seconds
(without alignments)
13680.175 Million cell updates/sec

Title: US-09-820-788A-3

Perfect score: 18731

Sequence: 1 agccttacaagtctgga.....ccagggtcagtcgaggt 10278

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p, model -DBV=xlp
-Q=/cgm2.1/USPTO.spool.p/US09820788/runat.24022004.141426.20037/app.query.fasta_1.12174
-DB=SPREMBL_25 -QFWT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09820788 @CGN 1.1 785 @runat.24022004.141426.20037 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp Vertebrate: *
14: sp_unclassified: *
15: sp_rvirus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1818	9.7	500 4	Q16753 homo sapien

	2	1678	9.0	497	6	Q8WNR5
	3	1655	8.8	497	6	Q865W1
	4	1461	7.8	497	4	Q16804
	5	1318.5	7.0	500	6	Q29454
	6	1311.5	7.0	500	6	Q9TUJ4
	7	1252.5	6.7	500	6	Q9TUJ5
	8	1234.5	6.6	500	11	Q9UKY7
	9	1234.5	6.6	500	11	Q9LW87
	10	1172	6.3	500	11	Q8VCX0
	11	1165.5	6.2	500	11	Q8CIM7
	12	1162.5	6.2	500	11	Q9DBJ5
	13	1137	6.1	504	11	Q921V1
	14	1131	6.0	504	11	Q84530
	15	1122	6.0	504	11	Q64529
	16	1101	5.9	504	11	Q8BVD2
	17	751.5	4.0	505	13	Q7SYW2
	18	629	3.4	5146	6	Q8SPM4
	19	577.5	3.1	4123	4	Q75851
	20	532.5	2.8	1953	5	Q9BIT7
	21	521	2.8	4998	11	Q8CG65
	22	519	2.8	5146	6	Q8SPM4
	23	513.5	2.8	1682	11	Q9QZR9
	24	509.5	2.7	1953	5	Q9BIT7
	25	507.5	2.7	496	13	Q803J0
	26	505	2.7	497	13	Q7ZU60
	27	495	2.6	1682	11	Q9QZR9
	28	494	2.6	498	13	Q9IAX8
	29	492.5	2.6	498	13	Q9PTR2
	30	491	2.6	501	11	Q9QXF7
	31	490	2.6	676	6	Q95JC9
	32	487.5	2.6	498	13	Q9PTR1
	33	484.5	2.6	13288	6	Q18758
	34	483	2.6	2284	5	Q9VPG1
	35	482.5	2.6	1691	11	Q9ESQ2
	36	479.5	2.6	1802	5	Q17163
	37	476.5	2.6	1869	11	Q9QZS0
	38	473	2.5	501	11	Q8BR78
	39	471.5	2.5	505	13	Q9IAT1
	40	470.5	2.5	502	11	Q924D1
	41	470.5	2.5	529	11	Q9CRO9
	42	469	2.5	2157	11	Q9Z1R1
	43	468.5	2.5	2158	11	Q7TSC1
	44	467.5	2.5	213	4	Q9NSM1
	45	467.5	2.5	1449	13	Q802B5

ALIGNMENTS

RESULT 1

ID	Q16753	PRELIMINARY;	PRT;	500 AA.
AC	Q16753;			
DT	01-NOV-1996 (TRENBLrel. 01, Created)			
DT	01-NOV-1996 (TRENBLrel. 01, Last sequence update)			
DT	01-JUN-2003 (TRENBLrel. 24, Last annotation update)			
DE	Debrisoquine 4-hydroxylase mutant allele (CYP2D6-M1)			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Gonzalez F.J.;			
RL	Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.			
CC	-!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.			
DR	EMBL; M33189; AAA35737.1; -			
DR	HSSP; P00179; 1DT6.			
DR	GO; GO:0004497; F:monooxygenase activity; IEA.			
DR	GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . . ; IEA.			
DR	GO; GO:0006118; P:electron transport; IEA.			
DR	InterPro; IPR001128; Cytochrome P450.			
DR	InterPro; IPR008069; EP450_CYP2D.			
DR	Pfam; PF00067; p450; 1.			

DR PRINTS; PRO1686; EP4501CVP2D.
DR PRINTS; PRO0385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 55889 MW; D5293B9BF74692C8 CRC64;

Alignment Scores:

Pred. No.: 1-55e-103 Length: 500
Score: 1818.00 Matches: 459
Percent Similarity: 34.38% Conservative: 12
Best Local Similarity: 33.52% Mismatches: 19
Query Match: 9.71% Indels: 900
DB: 4 Gaps: 8

US-09-820-788A-3 (1-10278) x Q16753 (1-500)

```
QY 2078 ATGGGGCTAGAGCACTGGTGCCTCGCCATGATAGTGGCCATCTTCTGCTCTGCTG 2137
Db 1 MetGlyLeuGluAlaLeuValProLeuAlaValAlaLeuValPheLeuLeuVal 20
QY 2138 GACCTGATGACCGGACCAAGCTGGGCTGACGCTACCGCCAGGTCCTCGCCACTG 2197
Db 21 AspLeuMetHisArgGlnArgTrpAlaAlaArgTyrSerProGlyProLeuProLeu 40
QY 2198 CCGGGCTGGGCAACTGCTGCATGTGACCTTCCAGAACACACCATCTGCTCGACCA 2257
Db 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspG1 60
QY 2258 GGTGAGGGAGGAGTCTTGGAGGGCGGACAGAGTCTCTGAGGATGCCACCACGACAA 2317
Db 60 n----- 60
QY 2318 CATGGTGTGGTTAAACCAAGCTGGATCAGAACGAGCTGAGAGGGGAAGCAGG 2377
Db 60----- 60
QY 2378 TTTGGGGACGTTCTTGGGGAAGACATTTATACATGTCATGAAGGACTGGATTTTCAA 2437
Db 60----- 60
QY 2438 AGGCCAAGGAAGATAGGCAAGGCTGGAGGTGGAGTGGACTTGGCAGTGGGCATGC 2497
Db 60----- 60
QY 2498 AAGCCATTGGCAACATATGTTATGGAGTACAAAGTCCCTCTCTGACACCAGAAGA 2557
Db 60----- 60
QY 2558 AAGCCCTTGGGAATGGAAGATGAGTTAGTCTCTGAGTGGCGTTTAAATCAGCAATCGAGG 2617
Db 60----- 60
QY 2618 ATGAAGGGGGTGCAGTGCACCGGTTTCAAACCTTTTGCACGTGGGTCTCTCGGGCTCACT 2677
Db 60----- 60
QY 2678 GCTACCGGCATGGACCATCATCTCGGGAATGGGATGCTAACTGGGGGCTCTCGGCAATTT 2737
Db 60----- 60
QY 2738 TGGTGACTCTTGAAGGTATACCTGGGTGAGCGCATCAAACCTGAGTTCTTCCATCACAG 2797
Db 60----- 60
QY 2798 AAGGTGTGACCCACCCCTGCCACGATCAGAGGCTGGGTCTCTCTTCCACCTGC 2857
Db 60----- 60
QY 2858 TCACCTCTGGTAGCCCGGGGGTCTGCCAAGGTTCAAATAGGACTAGGACCTGTAGTCTG 2917
Db 60----- 60
QY 2918 GGTGATCTCGGTTTGACAAGAGGCCCTGACCCCTCTCTGAGTTGCGGGCGCGCTTCG 2977
```

```
Db 61-----:::|||||
QY 2978 GGGACGTGTTTTCAGCTTGCAGCTGGCGCTGGAGCGCGGTGGTCTGCTCAATGGGCTGGCGG 3037
Db 66 1yAspValPheSerLeuGlnLeuAlaTrpThrProValValLeuAsnGlyLeuAlaA 86
QY 3038 CCGTGGCGGAGCGATGTTGACCCGCGGAGGACACGCGGACGCGCGCTGGCGCCA 3097
Db 86 1aValArgGluAlaMetValThrArgGlyGluAspThrAlaAspArgProValProI 106
QY 3098 TCTACCAAGTCTCGGCTTCCGGCGCGCTTCCCAAGCAAGCGCGGTGGGGGACAGAGA 3157
Db 106 1eThrGlnIleLeuGlyPheGlyProArgSerGlnGlyLysGln----- 120
QY 3158 CCGCTTTTCCGTGGGCCCCGGGTGGACAGTACCGTAGCCCAAGCAGCGCCGACAGGGCG 3217
Db 120----- 120
QY 3218 TGGGGTCTTGACGTGAAACAGAGATAAAGCCAGCGAGTGGGTGAGGACAGTGGGCCA 3277
Db 120----- 120
QY 3278 GGAACACCTGTCACGGGGAGGTGCGAGTCTGTGGGTGGGAGGGCGGGCTACTGC 3337
Db 120----- 120
QY 3338 CCAGACCCGCGAGAACGCCGCTGGCGGAGGCTGATGCGTGCAGTGGCGGTGGCGGGAC 3397
Db 120----- 120
QY 3398 CGCGCTATGCTGCGGGCTCAGTGTGGCGGGAAGGGGGGATCTTCTTCTGAGTGAAG 3457
Db 120----- 120
QY 3458 GTGGTCAAGGTGGCGAGAGACAGGTGGGGGCAAAACCCGCGCAGGCGGAGCAATG 3517
Db 120----- 120
QY 3518 TGGGTGAGCAAGAGTGGGCGCTGTGCCCCAGCTGAGCCGGGTAGGACTCGGGAGACC 3577
Db 120----- 120
QY 3578 TTGTGAGGCCAGGTTGGAGTGGGTGGCGGAGGTGGGCGCAAGGCTTTCATGGCAAC 3637
Db 120----- 120
QY 3638 GCCCAGCTGTCGTCGCGCCCGAGGGTGATCCTGTGCGCTATGGCCCGCGTGGCGC 3697
Db 121-----ArgGlyValProGlyAlaLeuTrpAlaArgValAlaA 133
QY 3698 GAGCAGAGGCGCTTCTCGTGTCCACTTGCACCTTGGCGCTTGGCGCAAGAGTCTGCTG 3757
Db 133 rgAlaGluAlaLeuLeuArgLeuHisLeuAlaGlnLeuGlyProGlyGlnGluValAlaG 153
QY 3758 GAGCAGTGGGTGACCGGAGGCGCTGCTCTTGTGCGCTT-CGCGCAACAAGCCGCT 3816
Db 153 1yAlaValIleSerAsnValIleAlaSerLeuThrCysGlyArgArgGlnProLeuArg- 172
QY 3817 GGGTGTGGGCAAGAGGCAAAAGCGGGAACCTGGGAAGGGCGGGGACGAGAGGCAAC 3876
Db 172----- 172
QY 3877 CCCTTACCCGCACTCTCCCAACCCCGAGAGCGCCCTTTCGCCCAAGCGCTTGGACA 3936
Db 173-----ArgProPheArgProAsnGlyLeuLeuAspL 183
QY 3937 AAGCGGTGAGCAACGTGATCGCTCCCTCACCTCGGGCGCGCTTTCAGTACACGACC 3996
Db 183 ysAlaValIleSerAsnValIleAlaSerLeuThrCysGlyArgArgGlnProLeuArg- 203
QY 3997 CTGCTTCTCAGGCTGCTGAGCTAGCTCAGGAGGACCTGAAGAGGAGTTCGGGCTTTC 4056
```


QY 6217 TCTGGTACCCCATCCCTACGAGCTTGTGCTGTGCCCGC 6259
 |||||
 Db 486 eluValThrProSerProTyrGluLeuCysAlaValProArg 500

RESULT 2

Q8WNR5 PRELIMINARY; PRT; 497 AA.
 ID Q8WNR5;
 AC Q8WNR5;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Cytochrome P450 2D.
 OS Macaca fuscata (Japanese macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Macaca.
 OX NCBI_TaxID=9542;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Hichiya H., Yamamoto S., Asaoka K., Narimatsu S.;
 RT "Molecular cloning and functional analysis of a Japanese monkey CYP2D enzyme."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC - 1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF301911; AAL73443.1; ..
 DR GO; GO:0004497; F:monooxygenase activity; IEA.
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR011128; Cytochrome P450.
 DR InterPro; IPR008069; EP450_CYP2D.
 DR Pfam; PF00667; P450; 1.
 DR PRINTS; PRO1686; P450I_CYP2D.
 DR PRINTS; PRO0385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 DR Heme; Monooxygenase; Oxidoreductase.
 SK SEQUENCE 497 AA; 55566 MW; 4C06EDD12F044D25 CRC64;

Alignment Scores:

Pred. No.: 6,44e-95 Length: 497
 Score: 1678.00 Matches: 445
 Percent Similarity: 33.00% Conservative: 17
 Best Local Similarity: 31.79% Mismatches: 35
 Query Match: 8.96% Indels: 903
 DB: 6 Gaps: 9

US-09-820-788a-3 (1-10278) x Q8WNR5 (1-497)

QY 2078 ATGGGGCTAGAGCACTGGTGGCCCTGGCCATGATAGTGGCCATCTTCTGCTCCTGGTG 2137
 |||||
 Db 1 MetGluLeuAspAlaLeuValProLeuAlaValThrValAlaIleLeuLeuVal 20
 QY 2138 GACCTGATGCACGGCACCACACGCTGGGTGCACGCTACCCGCCAGGTCCCTGCACTG 2197
 |||||
 Db 21 AspLeuMetHisArgGlnArgTrpAlaAlaArgTyrProGlyProLeuProLeu 40
 QY 2198 CCCGGCTGGGCAACCTGCTGCATGTGGACTTCAGACTTCAGAACACCATCTGTCGACCA 2257
 |||||
 Db 41 ProPheLeuGlyLys-LeuLeuHisValAspPheGlnAenThrProTyrCysPheAspG1 60
 QY 2258 GGTGAGGGAGGAGTCTCTGGAGGGCGGAGAGGTCTTGAGGATGCCCCACCACCAACAA 2317
 |||||
 Db 60 n----- 60
 QY 2318 CATGGTGTGGTTAAACACACAGGCTGGATCAGAAGCAGGCTGAGAAGGGGAAGCAGG 2377
 |||||
 Db 60 ----- 60
 QY 2378 TTTGGGGGACGTTCTCTGGGGAAGGACATTTATACATGCGATGAAGGACTGGATTTTCCAA 2437
 |||||
 Db 60 ----- 60
 QY 2438 AGGCCAAGGAAGTAGGCAAGGGCCTGGAGGTGGAGCTGGACTTGGCAGTGGGCATGC 2497

Db 60 ----- 60
 QY 2498 AAGCCCATTTGGGCAACATATGTTATGGAGTACAAAAGTCCCTTTCTGCTGACACCCAGAAGGA 2557
 |||||
 Db 60 ----- 60
 QY 2558 AAGCCCTTGGGAATGGAAGATGAGTTAGTCTGAGTGCCGTTTAAATCACGAATCGAGG 2617
 |||||
 Db 60 ----- 60
 QY 2618 ATGAAGGGGGTGTCAGTGACCCGGTTCAAACTTTTGCACTGTGGGTCTCTCGGGCTCACT 2677
 |||||
 Db 60 ----- 60
 QY 2678 GCTCACCGGCATGACCATCATCTGGGAATGGGATGTAACTGGGGCTCTCGGCAATTT 2737
 |||||
 Db 60 ----- 60
 QY 2738 TGGTGACTCTTGCACAGGTTCATACCTGGGTGAGGCATCCAAACTGAGTTCTCTCCATCACAG 2797
 |||||
 Db 60 ----- 60
 QY 2798 AAGGTGTACCCCAACCCCTGCTCCACGATCAGGAGGTGGGTCTCTCTCTTCCACCTGC 2857
 |||||
 Db 60 ----- 60
 QY 2858 TCATCTCTGTGTAGCCCGGGGGTCTCAAGGTTCAAATAGGACTAGGACCTGAGTCTG 2917
 |||||
 Db 60 ----- 60
 QY 2918 GGGTGATCCTGGCTTGACACAGAGCCCTGACCCCTCTCTCTGAGTTGGGGCGCGCTCG 2977
 |||||
 Db 61 ----- 66
 QY 2978 GGGACGTGTTTCAGCTCTGAGCTGGCTGGACCGCGGTGGTCTGCTCAATGGGTGGCGG 3037
 |||||
 Db 66 lyAspValPheSerLeuGlnLeuAlaTrpThrProValValLeuAsnGlyLeuAla 86
 QY 3038 CCGTGCCGAGGCGATGTTGACCCGCGCGAGGACACGGCGGACACGGCGGCTCGCGCCA 3097
 |||||
 Db 86 laValAlaGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProValProI 106
 QY 3098 TCTACAGCTCTTGGGCTTGGGGCGGTTCCCAAGGCACAGCGGGTGGGGGACAGAGA 3157
 |||||
 Db 106 leThrGlnIleLeuGlyPheGlyProArgSer-Gln----- 117
 QY 3158 CCGGCTTTCTCGTGGGCGCCCGGGTGGACAGTACCGTAGCCCAAGCAGCGCCGACAGGGCG 3217
 |||||
 Db 117 ----- 117
 QY 3218 TGGGGTCTGGACCTGAAACAGAGATAAGCCAGCGAGTGGGTGAGGACAGTGGGCCA 3277
 |||||
 Db 117 ----- 117
 QY 3278 GGAACACCTGCACGGGGAGGTCCGAGTCTGTGGGTGGGAGGGGGGGGCTACTGC 3337
 |||||
 Db 117 ----- 117
 QY 3338 CCAGACCGCCAGAAAGCCCGGTGGCGAGGCTGATGCTCGAAGTGGCGGTGGCGGGGAC 3397
 |||||
 Db 117 ----- 117
 QY 3398 CGCGCTTATGCTCGGGCTCAGTGTGGCGGGACGGCGGGATCTTCTTGAAGTGAAG 3457
 |||||
 Db 117 ----- 117
 QY 3458 GTGGTCAGGTGGGCAGAGCAGGTGGGGCCAAACCCCGCCCGGAGGAGCAATG 3517
 |||||
 Db 117 ----- 117
 QY 3518 TGGGTGAGCAAGAGTGGGCCCTCTGTCCAGCTGGACCGGGCTAGGACTGCGGGAGACC 3577

Db 117 ----- 117
QY 3578 TTGTGGAGCCAGGGTTGGAGTGGGTGGCGGAGGGTGGGGCCCAAGGCCCTTCATGGCAAC 3637
Db 117 ----- 117
QY 3638 GCCCAGCTGTCCGTCCGGCCCGCCAGGGGTGATCCTGTCTGGGCTATGGGCCCGCGTGGCGC 3697
Db 118 ----- 118
QY 3698 GAGCAGAGCGCTTCTCCGTGTCCACTTGGCAACTTGGGCTGGGCAAGAGTCGCTG 3757
Db 130 GluGlnArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysSerGly 149
QY 3758 GAGCACTGGGTGACCGAGGAGCGCGCTGCTTGTGCGGCTTTCGCGACCAAGCCGGTG 3817
Db 150 ThrGlnTrpValThrGlyGluAlaLaCysLeuCysAlaPheAlaAsn----- 166
QY 3818 GGTGATGGCAGNAGGGCACAAGCCGGAACTGGGAAGCGGGGGAGCGAGAGGCAACC 3877
Db 166 ----- 166
QY 3878 CCTTACCCGATCTCCACCCCGCCAGGAGCGCCCTTTCGCCCAAGCGCTCTTGACAA 3937
Db 167 -----His-Ser-----GlyArgProPheArgProAsnGlyLeuLeuAspLys 180
QY 3938 AGCCGTGAGCAACGTGATCGCTCCCTCACCTCGCGGCGCGCTTCGAGTACGACGCC 3997
Db 180 alaValSerAsnValIlealaSerLeuThrCysGlyArgArgPheGluTyAspAspPr 200
QY 3998 TCGTTCCTCAGGCTCTCGACTAGCTCAGAGGGACTGAAGGAGGAGTCCGGCTTTCT 4057
Db 200 oArgPheLeuArgLeuLeuAspLeuAlaGlnGlyLeuLysGluLuproGlyPheLe 220
QY 4058 GCGCGAGGTGCGGAGCGAGAGCCGAGGAGTCTCTCGAGGCGGAGTCTCTGAGAGTGCC 4117
Db 220 uArgGlu----- 222
QY 4118 GGGGTGGACTGGGGCTCCGAAGGCGAGGATTTGCATAGATGGGTTTGGGAAAGGACAT 4177
Db 222 ----- 222
QY 4178 TCCAGAGAGCCCACTGTGAAGAGGCGCTGGAGGAGGAGGACATCTCAGACATGGTGG 4237
Db 222 ----- 222
QY 4238 TGGGAGAGTGTGCCGGGTGAGGGGACACAGAGAGGCCAAGGACTCTGTACCCCGT 4297
Db 222 ----- 222
QY 4298 CCACGTTGAGATTTTCGATTTTAGGTTTCTCTCTGGGCAAGGAGAGGGTGGAGGCTG 4357
Db 222 ----- 222
QY 4358 GCACTTGGGAGGAGCTTGGTGAAGTCAAGTGAAGACAGGAGCGCCCTGGGTCTACCT 4417
Db 222 ----- 222
QY 4418 GGAGATGGTGGGGCTGAGACTTGTCCAGGTGAACGCGAGACACAGAGGGGATTGAGAC 4477
Db 222 ----- 222
QY 4478 CCCGTTCTGTCTGGTGTAGTGTCTGAATGCTGCTCCCGTCTCCCTGCGACATCCAGCGCT 4537
Db 223 -----ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe 236
QY 4538 GCGTGGCAAGGTCTACGCTTCCAAAAGGCTTTCCTGACCCAGCTGGATGAGCTCCTAAC 4597
Db 236 uAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGACACAGGATGACTGGGACCCAGCCAGCCACCCGAGACCTGACTGAGGCGCTTCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276

QY 4658 GGCAAAGAGAGAGAGGTGAGAGTGGCTGCCAGCTGGGGGGCAAGGGTGGTGGGTTGAA 4717
Db 276 uAlaGluMetGluLys----- 281
QY 4718 COTCCAGAGGAGATGAGGGGAGGCTGGGCAAAAGGTTGGACCGAGTGCATCACCCGGCGA 4777
Db 281 ----- 281
QY 4778 GCCGCATCTGGGCTGACAGGTGCAGNATTGGAGGTCAATTTGGGGGCTACCCCGTTCTATC 4837
Db 281 ----- 281
QY 4838 CCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCAATGA 4897
Db 282 -----AlaLysGlyAsnProGluSerSerPheAsnAs 292
QY 4898 TGAGAACCTGCCCATAGTGGTGGGTAACTGTTCCTTGGCGGATGGTGCACCACTCGAC 4957
Db 292 pGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetValThrSerTh 312
QY 4958 CACGCTGGCGCTCGGCGCTCTGCTCATGATCCTACCTGGATGTGCAGCGTGAGCCCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuMetIleLeuHisProAspValGln----- 328
QY 5018 CTGGGGCCCAAGCGAGGACTGAGGGAGGAGGTACAGCTGGGGGCCCTCGGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCCGGGGCTTCCAGCACAGCGGTGGCCAGGCTCTGTAAAGCCTAACTTCCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGAGGAGAGAGAGTGTCCCTGGGTGCTGACCCATTTGTGGGGACGCGATGCTGT 5197
Db 328 ----- 328
QY 5198 CCAGTCCGTGTCCAAAGAGAGATCGACGAGTGTAGGACAGGTGGCGACACAGAGATG 5257
Db 329 -Arg-ArgValGlnGlnGluIleAspAspValIleGlyAlaValTrpArgProAlaLeu 347
QY 5258 GGTGACAGGCTCACATGCCCTACACCACTGCGTGTATTCACGAGGTGACGCGCTTGGG 5317
Db 348 GlyGluGlnAlaHisMetProTyThrProAlaValLysHisGluValGlnArgPheGly 367
QY 5318 GACATCATCCCGCTGAGTGTGACCCATATGACATCCCGTGCATCGAAGTACAGGCTTC 5377
Db 368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CGCATCCCTAAGGTAGGCTGGGCGCTCTCACCCAGCTCAGCACACGACCTGGTGA 5437
Db 388 ArgIleProLys----- 391
QY 5438 TAGCCCCAGCATGGCTACTGCCAGGTGGGCCCACTTAGGAACCTTGCCACCTAGTCTCT 5497
Db 391 ----- 391
QY 5498 CAATGCCACACACTGACTGTCTCCCACTTGGTGGGGGGTCCAGAGTATAGGACGGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTCCATCCAGAGCCCCGCTAGTGGGGGAGACNAACAGGACCTGCCAGATGTTG 5617
Db 391 ----- 391
QY 5618 GAGGACCCAGCGCTGACAGGAGGGGGGAGTGTGGGTGCTCTGAGAGGTGTGACTGC 5677
Db 391 ----- 391
QY 5678 GCCCTGTGTGGGTGCGAGAGGGGTACTGTGGAGCTTCTCGGGGCCGACGACTAGTTGACA 5737
Db 391 ----- 391

QY	5258	GGTGACCGCTCACATCCCTACACCACTGCGGTGATTTCACGAGTGCGAGCTTTGGG	5317	DT	01-NOV-1996	(TREMBLrel. 01, Created)
Db	348	GlyAspGluAlaHisMetProTyrThrThrAlaAlaIleHisGluValGlnArgPheGly	367	DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)
QY	5318	GACATCATCCCTCAGTGTGACCCATATGACATCCGTGACATCAAGTACAGGCTTC	5377	DE	01-JUN-2003	(TREMBLrel. 24, Last annotation update)
Db	368	AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe	387	OS	Cytochrome P450db1 (Fragment)	
QY	5378	CGATCCCTAAGTAGGCTGGCGCCCTCTCACCCAGCTCAGCACACCACTGGTGA	5437	OC	Homo sapiens (Human)	
Db	388	ArgIleProLys-----	391	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
QY	5438	TAGCCCCAGCATGGCTACTGCCAGGTGGGCCCACTTAGAACCCTGGCCACTAGTCCT	5497	OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.	
Db	391	-----	391	OX	NCBI_TaxID=9606;	
QY	5498	CAATGCCACCACTGACTGTCCCACTTGGGTGGGGGTCCAGAGTATAGGCAGGGCTG	5557	RN	[1]	
Db	391	-----	391	RP	SEQUENCE FROM N.A.	
QY	5558	GCCTGTCCATCCAGAGCCCCCGTCTAGTGGGAGACAAACAGGACCTGCCAAGATTGTG	5617	RC	TISSUE=Liver;	
Db	391	-----	391	RX	MEDLINE=89155788; PubMed=2466049;	
QY	5618	GAGGACCCAGCGCTCAGGAGAGGGGCGAGTGTGGTGCCTCTGAGAGGTGTGACTGC	5677	RA	Manns M.P., Johnson B.F., Griffin K.J., Tan E.M., Sullivan K.F.;	
Db	391	-----	391	RT	"Major antigen of liver kidney microsomal autoantibodies in idiopathic	
QY	5678	GCCCTGCTGTGGGTTCGAGAGGGTACTGTGAGCTTCTCGGCGCAGGACTAGTTGACA	5737	RT	autoimmune hepatitis is cytochrome P450db1.";	
Db	391	-----	391	RL	J. Clin. Invest. 83:1066-1072(1989).	
QY	5738	GAGTCCAGTGTGTGCCAGGCACTGTGTCTCCCGTGTGTGTGGTGGAGGGTCCCAG	5797	CC	-/- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.	
Db	391	-----	391	DR	EMBL; M24499; AAA36403.1; .	
QY	5798	CATCTAGAGTCCAGTCCCACTCTCACCTGCATCTCTGCCAGGAGACCACTCAT	5857	DR	HSSP; P00179; 1DT6.	
Db	392	-----	392	DR	GO; GO:0004497; F:monoxygenase activity; IEA.	
QY	5858	CACCAACTGTCTATCGTGTGAAGGATGAGGCGCTCTGGAGAGCCCTTCGGCTTCCA	5917	DR	GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . . ; IEA.	
Db	396	ethrAsnAsgProSerValLeuIysAspGluAlaValTrpGluLysProPheArgPheHi	416	DR	GO; GO:0006118; P:electron transport; IEA.	
QY	5918	CCCCAACACTTCTCGATGCCAGGCCACTTTGTGAAGCCGAGGCTTCTCGCTTT	5977	DR	InterPro; IPR001128; Cytochrome P450.	
Db	416	sProGluHisPheLeuAspAlaGlnGlyArgPheValLysProAspAlaPheLeuProH	436	DR	InterPro; IPR008069; EP450_CYP2D.	
QY	5978	CTCAGCAGTGTCTGTGGGAGCCCGCTCCCTGTCTCCCTTCCGTGGAGTCTTGACGGG	6037	DR	Pfam; PF00067; P450; 1.	
Db	436	e--Ser-----	437	DR	PRINTS; PRO1686; EP450ICYP2D.	
QY	6038	TATACCCAGAGCCAGGCTCACTGACGCCCTCCCTCCACAGCGCGCTGCATGC	6097	DR	PRINTS; PRO0385; P450.	
Db	438	-----	443	DR	PROSITE; PS00086; CYTOCHROME P450; 1.	
QY	6098	CTCGGGAGCCCTGGCCGCAATGAGCTTCTCTTTTACCTCCCTCTGTCAGCAC	6157	KW	Heme; Monooxygenase; Oxidoreductase.	
Db	444	LeuGlyGluProArgAlaArgMetGluPhePhePheThrCysLeuLeuGlnHis	463	FT	NON_TER	
QY	6158	TTTCAGTCTTCTCGTGGCCCGGACAGCCCGCCGAGGCACTCTCTGTCGTCAGCTTT	6217	SQ	SEQUENCE 373 AA; 42005 MW; 9FDED67B0BA487A4 CRC64;	
Db	464	PheSerPheSerValProThrGlyGlnProArgProSerHisGlyValPheAlaPhe	483	Alignment Scores:		
QY	6218	CTGGTACCCATCCCTACAGCTTTGTGTGTGTCGCCCGC 6259		Pred. No.:	1.4e-81	Length: 373
Db	484	LeuValSerProSerProTyrGluLeuCysAlaValProArg 497		Score:	1461.00	Matches: 353
RESULT 4				Percent Similarity:	41.60%	Conservative: 6
Q16804				Best Local Similarity:	40.90%	Mismatches: 14
ID	Q16804	PRELIMINARY; PRT; 373 AA.		Query Match:	7.80%	Indels: 490
AC	Q16804;			DB:	4	Gaps: 7


```
Qy 4163 TTTGGGAAAGGACATTCCAGGAGACCCCACTGTGAAGAGCGCTGGAGGAGGGGACA 4222
Db 98 ----- 98
Qy 4223 TCTCAGACATGGTCTGGGAGAGGTGTGCCGGGTCAAGGGGACACAGGAGAGGCCAAGG 4282
Db 98 ----- 98
Qy 4283 ACTCTGTACCCCGCTCCACGTTGGAGATTTCGATTTTAGTGTTCCTCTCTCTGGGCAAGGAG 4342
Db 98 ----- 98
Qy 4343 AGAGGGTGGAGGCTGGCACTTGGGGAGGGCACTTGGTGAGGTCAAGTGAAGGACAGGAG 4402
Db 98 ----- 98
Qy 4403 GCCCTGGGTCTACCTGGAGATGGCTGGGGCTGAGACTTGTCCAGGTGAAGCAGAGCAC 4462
Db 98 ----- 98
Qy 4463 AGGAGGATTGAGACCCCGTTCTGTCTGGTGTAGGTCTGAATGCTGTCCCGCTCTCCCT 4522
Db 99 -----ValLeuAsnAlaValProValLeuLe 107
Qy 4523 GCACATCCCAAGCGCTGGCTGGCAAGTCTACGCTTCCAAAGGCTTTCCTGACCCAGCT 4582
Db 107 uHisIleProAlaLeuAlaGlyIysValLeuArgPheGlnLysAlaPheLeuThrGlnLe 127
Qy 4583 GGATGACTCTTAAGTACAGCACAGGATGACCTGGGACCCAGCCAGCACCCCGAGACCT 4642
Db 127 uAspGluLeuLeuThrGluHisArgMetThrTrpAspProAlaGlnProProArgAspLe 147
Qy 4643 GACTGAGCGCTTCTCTGCAAGAGAGGAGGTGAGAGTGGCTGCCACGCTGGGGGCA 4702
Db 147 uThrGluAlaPheLeuAlaGluMetGluLys----- 157
Qy 4703 GGGTGGTGGTGAACCTCCAGGAGGAATGAGGGGAGGTGGGGCAAAAGTTGGACCA 4762
Db 157 ----- 157
Qy 4763 TGCATCACCGGCGAGCGGATCTGGGCTGACAGGTGCAGAAATGGAGGTCAATTGGGGG 4822
Db 157 ----- 157
Qy 4823 CTACCCCGTTCTATCCCTGAGTATCTCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGA 4882
Db 158 -----AlaLysGlyAsnProGle 163
Qy 4883 GAGCAGCTTCAATGATGAGAACTTGCATAGTGGTGGTAACTGTTCCTTCCCGGGAT 4942
Db 163 uSerSerPheAsnAspGluAsnLeuArgIleValAlaAspLeuPheSerAlaGlyMe 183
Qy 4943 GGTGACCACTCGACCAAGTGGCTGGGCTCTCTCTCATGATCTACACCTGGATGT 5002
Db 183 tValThrThrSerThrThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspVa 203
Qy 5003 GCACGCTGAGCCAGCTGGGGCCCAAGCAGGAGGACTGAGGAGGAGGTTACAGCTGGG 5062
Db 203 lGln----- 204
Qy 5063 GCCCTGGGCTTAGCTGGGACACCCGGGGCTTCCAGCACAGGGGTGGCGAGGCTCTGTGA 5122
Db 204 ----- 204
Qy 5123 AGCCTAACTCTTCCAAACACAGGAGGAAGAGAGTGTCCCTGGGTGCTGACCCATTGTG 5182
Db 204 ----- 204
Qy 5183 GGGACGCACTGTCTGTCCAGTCCGTGTCACACAGGAGATCCAGCAGCTGATAGGGCAGGTG 5242
Db 205 -----Arg--ArgValGlnGlnIleAspAspValIleGlyGlnVal 218
```

```
Qy 5243 CGGCGACCAAGAGATGGGTGACCAAGGTTCACATGCCCTACACCACTGCGGTGATTACAGAG 5302
Db 219 ArgArgProGluMetGlyAspGlnAlaHisMetProFyrThrThrAlaValIleHisGlu 238
Qy 5303 GTCACAGCGCTTTGGGACATCATCCCTCAGTGTGACCCATATGACATCCCGTGACATC 5362
Db 239 ValGlnArgPheGlyAspIleValProLeuGlyValThrHisMetThrSerArgAspIle 258
Qy 5363 GAAATGACAGGCTTCCGCACTCCCTAAAGTAGGCTGGCGCCCTCCTCACCCAGCTCAGC 5422
Db 259 GluValGlnGlyPheArgIleProLys----- 267
Qy 5423 ACCAGCACCTGGTGTAGATAGCCCCAGCATGGCTACTGCCAGGTGGGCCCACTCTAGGAACCC 5482
Db 267 ----- 267
Qy 5483 TGGCCACCTAGTCTCTCAATGCCACCACTGACTGTCCCACTTGGGTGGGGGTCCAGA 5542
Db 267 ----- 267
Qy 5543 GTATAGCAGGCTGGCTGTCCATCCAGAGCCCCGTCTAGTGGGGAGACAAACAGGA 5602
Db 267 ----- 267
Qy 5603 CCTGCCAGAATTTGGAGGACCCAGCGCTTCAGGGAGAGGGGCGAGTGTGGGTGCCTCT 5662
Db 267 ----- 267
Qy 5663 GAGAGTGTGACTGCGCCCTGTCTGTGGGTTCGGAGAGGTTACTGTGGAGCTTCTCGGGCG 5722
Db 267 ----- 267
Qy 5723 CAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAGGCAGTGTGTCTCCCGTGTGTTGG 5782
Db 267 ----- 267
Qy 5783 TGGCAGGGTCCAGCATCTTAGAGTCCAGTCCCACTCTCACCTTGCATCTCTCTGCCCA 5842
Db 267 ----- 267
Qy 5843 GGGAAACGACACTCATCACCACTGTCTCATCGTGTCTGAAGGATGAGGCCCTCTGGGAGAA 5902
Db 268 Gly-ThrThrLeuIleThrAsnLeuSerValLeuLeuAspGluAlaValTrpGluLys 287
Qy 5903 GCCCTTCCGCTCCACCCGAAACACTTCTTGGATGCCAGGGCCACTTCTGTAAGCCGGA 5962
Db 287 sProPheArgPheHisProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGle 307
Qy 5963 GGCCTTCTCTGCTTCTCTCAGCAGGTCCTGTGGGGAGCCCGCTCTCTGCTCCCTTCCGT 6022
Db 307 uAlaPheLeuProPhe--Ser----- 313
Qy 6023 GGAGTCTTGAGGGGTATCACCCAGGAGCGAGCTCACTGACGCCCTCTCCCTCCCCACA 6082
Db 314 -----Ala 314
Qy 6083 GGGCCGCTGCATGCTCTCGGGAGCCCTGGCCGCGCATGGAGCTCTCTCTTCTTTCACC 6142
Db 315 GlyArgArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThr 334
Qy 6143 TCCTCTGTCGACACTTTCAGCTTCTCTCGTGGCCGCGGAGCAGCCCGGGCCAGCACCTCT 6202
Db 335 SerLeuLeuGlnHisPheSerPheSerValProThrGlyGlnProArgProSerHisHis 354
Qy 6203 CGTGTGTCAGCTTCTTGTGACCCCATCCCTTACAGCTTTGTGTGTGTGTGTGTGTGTGTGT 6259
Db 355 GlyValPheAlaPheLeuValThrProSerProFyrGluLeuCysAlaValProArg 373
RESULT 5
Q29454 ID PRELIMINARY; PRT; 500 AA.
AC Q29454;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
```

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome P-450 IID.
GN CYP2D.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=PBVL 180; TISSUE=Liver;
RX MEDLINE=93011103; PubMed=139678;
RA Tsunooka Y., Matsuo Y., Higuchi R., Ichikawa Y.;
RT 'Characterization of the cytochrome P-450IID subfamily in bovine
liver. Nucleotide sequences and microheterogeneity.';
RL Eur. J. Biochem. 208:739-746(1992).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; X68481; CAA48501.1; -;
DR PIR; S37284; S37284.
DR HSP; P00179; IDT6.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008089; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450_CYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 55921 MW; 4F62F39050E2BED6 CRC64;

Alignment Scores:

Pred. No.: 8,75e-73 Length: 500
Score: 1318.50 Matches: 379
Percent Similarity: 30.36% Conservative: 47
Best Local Similarity: 27.01% Mismatches: 71
Query Match: 7.04% Indels: 906
DB: 6 Gaps: 12

US-09-820-788a-3 (1-10278) x Q29454 (1-500)

QY 2078 ATGGGGCTA-----GAAGCATGGTGGCCCTGGCCCATGATAGTGGCCATCTTCCTG 2128
DB 1 MetGlyLeuLeuSerGlyAspThrLeuGlyProLeuAlaValAlaLeuLeuPheLeu 20
QY 2129 CTCCTGGTGGACTGATGACCGCACCAACGCTGGCTGCAGCTACCCGCGAGTCCC 2188
DB 21 LeuLeuLeuAspLeuMethi.sargArgSerA.gtrpAlaProArgTyrProGlyPro 40
QY 2189 CTGCCACTCCCGGGCTGGGCAACCTTGTGTCATGTGGACTTCCAGAACACACCATCTG 2248
DB 41 ThrProLeuProValLeuGlyAsn-LeuLeuGlnValAspPhe----- 54
QY 2249 CTTGACACAGGTGAGGAGAGAGTCTCGAGGGCGGCAGAGGTCTGAGGATGCCCCACC 2308
DB 55 -----GluAsp-Pro----- 57
QY 2309 ACCAGCAACATGGGTGGTGGGTTAAACACACAGGCTGGATCAGAGCCAGGCTGAGAAGG 2368
DB 57 ----- 57
QY 2369 GGAAGCAGGTTTGGGGGACGTTCTCTGGGAAGGACATTTATATCATGGCATGAAGGACTGG 2428
DB 57 ----- 57
QY 2429 ATTTTCAAAGGCCAAGGAGAGTAGGCGAAGGGCCTGGAGGTGGAGCTGGACTTGGCAG 2488
DB 57 ----- 57
QY 2489 TGGGCATGAAGCCCATTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACA 2548

Db 57 ----- 57
QY 2549 CCAGAGAAAGCCCTTGGGAATGGAAGATGAGTTAGTCTCTGAGTGCCTTTAATCAGC 2608
Db 57 ----- 57
QY 2609 AAATCGAGGATGAAGGGGGTGCGATGACCGGTTCAAACCTTTTGCACTGTGGGTCTCTCG 2668
Db 57 ----- 57
QY 2669 GGCCTCACTGCTACCGGCATGGACCATCATCTGGGAATGGGATGCTAACTGGGGCCTCT 2728
Db 57 ----- 57
QY 2729 CGGCAATTTTGGTGACTCTTCAAGGTCACTACCTGGGTGACGCATCCAAACTGAGTTCT 2788
Db 57 ----- 57
QY 2789 CCATCAGAAAGTGTGACCCCAACCCCTGCCCCACGATCAGAGGCTGGGTCTCTCTCT 2848
Db 57 ----- 57
QY 2849 TCCACTCTCACTCTCTGCTAGCCCCGGGGTCTGTCAAAGTTCAATAGGACTAGGACC 2908
Db 58 -----ArgProSerPheAsn----- 62
QY 2909 TGTAGTCTGGGGTGATCTCTGGCTTGACAAGAGGCCCTGACCTCCCTCTGCAGTTGCGGC 2968
Db 63 -----GlnLeuA.rga 66
QY 2969 GCCCTTCGGGACGTTTTCAGCTGACGCTGGCTGACGCGCGGTGTGTCGTCTCAATG 3028
Db 66 rArgPheGlyAsnValPheSerLeuGlnValTrpThrProValValValLeuAAsnG 86
QY 3029 GGCCTGGCGCGCTGCGGAGCGATGTGACCCCGCGCGAGACACGCGCCGACCGCCCGC 3088
Db 86 lyLeuAlaAlaValArgGluAlaLeuValTyrArgSerGlnAspThrAlaAspArgProp 106
QY 3089 CTGCGCCCATCATCACAGGTCTCTGGCTTGGCGTGGCGCGGTTCCTCCAGGCAAGCGCGGTGGG 3148
Db 106 roProAlaValTyrGluHisLeuGlyTyrGlyProArg----- 118
QY 3149 GGACAGAGACCGCGTTTCCGTGGGCGCCCGGTGGACAGTGCCTGAGCCAGCAGCGCC 3208
Db 118 ----- 118
QY 3209 CACAGGCGTGGGGTCTCTGGACGTGAACAGAGATAAAGCCAGCGAGTGGGCTGAGGAC 3268
Db 118 ----- 118
QY 3269 AGTGGGCCAGGAACACCTGCA CGGGGGAGGTGGAGTCTGTGGGTGGGAGGGGGCGG 3328
Db 118 ----- 118
QY 3329 GGCTACTGCCAGACCCCGCAGAACCCGCTGGGGAGGCTGATCGTCGTAAGTGGCGGT 3388
Db 118 ----- 118
QY 3389 GCGGGGACCGCGCTATGCTCGGGCTCAGTGTGGGGGACGCGGCGGGATCTTCTTTG 3448
Db 118 ----- 118
QY 3449 AGTGGAAAGGTGGTCAAGGTGGGCGAGACGAGGTGGGCGCCNAACCCCGCCAGGCGAG 3508
Db 118 ----- 118
QY 3509 GGAGCAATGTGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGCTAGGAGCTG 3568
Db 118 ----- 118
QY 3569 CCGGAGACCTTGTGAGCGCCAGGCTTGGAGTGGGTGGCGGAGGTGGGCGCAAGGCCTT 3628
Db 118 ----- 118

Db	395	-----		-----	Gly-Th	396
QY	5849	GACACTCATCAACACTGCTCATCGTGTGATGAGATGAGCGCGTGGGAGAGCCCTT		-----	Gly-Th	5908
Db	396	rThrLeuIleThrAsnLeuSerSerValLeuIysAspGluThrValrPgluLysProPh		-----	Gly-Th	416
QY	5909	CGCTTCCACCCCGAACACTTCTCGATGCCAGGCGCACTTGTGAAGCCGAGGCTT		-----	Gly-Th	5968
Db	416	eargPheHisProGluHisPheLeuAspAlaGlnGlyArgPheValLysGlnGluAlaPh		-----	Gly-Th	436
QY	5969	CTGCGCTTCTCAGCAGGTGCTGTGGGAGCGCGCTGCTCCCTTCCGTTGGAGTC		-----	Gly-Th	6028
Db	436	eileProPhe-Ser		-----	Gly-Th	440
QY	6029	TTGAGGGGTATACCCAGGAGCCAGGCTCATGTAGCGCCCTCCCTCCCCACAGCGCGC		-----	Gly-Th	6088
Db	441	-----		-----	Gly-Th	443
QY	6089	CGTGCATGCTCGGGAGCCCTGGCGCGCATGGAGCTTCTTCTTCTTCACTCCCTG		-----	Gly-Th	6148
Db	444	ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeu		-----	Gly-Th	463
QY	6149	CTGCAGCACTTCACTTCTCGTGGCGCGCGAGCGCCCGCCAGCCACTTCTGTGTC		-----	Gly-Th	6208
Db	464	LeuGlnHisPheSerPheSerValProAlaGlyGlnProArgProSerGluHisGlyVal		-----	Gly-Th	483
QY	6209	GTCAGCTTCTGTGACCCATCCCTACGAGCTTGTGTGCGCCCGC		-----	Gly-Th	6259
Db	484	PheAlaPheLeuValThrProAlaProTyrGlnLeuCysAlaValProArg		-----	Gly-Th	500
RESULT 6						
Q9TUJ4	PRELIMINARY;	PRT;	500	AA.		
AC	Q9TUJ4;					
DT	01-MAY-2000 (TrEMBLrel. 13, Created)					
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)					
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)					
DE	Cytochrome P450 2B11.					
OS	Oryctolagus cuniculus (Rabbit).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.					
OX	NCBI_TaxID=9986;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=New Zealand White; TISSUE=Liver;					
RX	MEDLINE=98391821; PubMed=9722658;					
RT	Yamamoto Y., Ishizuka M., Takada A., Fujita S.;					
RT	"Cloning, tissue distribution, and functional expression of two novel					
RT	rabbit cytochrome P450 isozymes, CYP2D23 and CYP2D24.";					
RL	J. Biochem. 124:503-508(1998).					
CC	-/- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.					
DR	EMBL; AB008785; BAA84473.1; --					
DR	PIR; JE0258; JE0258.					
DR	HSP; P00179; 1DT6.					
DR	GO; GO:004497; F:monooxygenase activity; IEA.					
DR	GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.					
DR	GO; GO:0006118; P:electron transport; IEA.					
DR	InterPro; IPR001128; Cytochrome P450.					
DR	InterPro; IPR008069; EP450_CYP2D.					
DR	Pfam; PF00067; P450; 1.					
DR	PRINTS; PR01686; EP450_CYP2D.					
DR	PRINTS; PR00385; P450.					
DR	PROSITE; PS00086; CYTOCHROME P450; 1.					
KW	Heme; Monooxygenase; Oxidoreductase.					
SQ	SEQUENCE 500 AA; 55604 MW; 434EC7C86BF6305B CRC64;					
Alignment Scores:						
Pred. No.:	2,36e-72	Length:	500			
Score:	1311.50	Matches:	379			
Percent Similarity:	29.58%	Conservative:	36			
Best Local Similarity:	27.01%	Mismatches:	82			
Query Match:	7.00%	Indels:	906			

DB:	6	Gaps:	9
US-09-820-788A-3 (1-10278) x Q9TUJ4 (1-500)			
QY	2078	ATGGGGCTA-----GAAGCACTGGTCCCTGGCCATGATAGTGGCCATCTCTCTG	2128
Db	1	MetGlyLeuLeuSerGlyGluAlaLeuAlaProLeuAlaValAlaValAlaLeu	20
QY	2129	CTCTGGTGGACCTGATGACCGGACCAACGCTGGGTGGACGCTACCGCCAGGTCC	2188
Db	21	LeuLeuValAspLeuMetHisLysArgProArgTrpAlaAlaArgTyrProProGlyPro	40
QY	2189	CTGCCACTGCGCGGCTGGCAACTTGTGATGTGACTTCCAGAACACACACTACTG	2248
Db	41	ValGlyIleProGlyLeuGlyAsn-LeuLeuGlnValAspPheArgGlyIleProAsnCy	60
QY	2249	CTTCGACCAAGTGGAGGAGGCTCTGGAGGGCGGAGAGGTCTGAGGATGCCCCACC	2308
Db	60	sPhe-Arg-----	62
QY	2309	ACCAGCAAAACATGGGTGGTGAACCAACAGGCTGGATCAGAACCCAGGCTGAGAAG	2368
Db	62	-----	62
QY	2369	GGAGCAGGTTTGGGGGACGTTCTCTGGGGAAGACATTTATACATGGCATGAAGGACTG	2428
Db	62	-----	62
QY	2429	ATTTCCAAAGGCCAAGAGAGTAGGGCAAGGGCTGGAGGTGGAGCTTGGCAG	2488
Db	62	-----	62
QY	2489	TGGGATGCAAGCCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTCTCTGTGACA	2548
Db	62	-----	62
QY	2549	CCAGAAGGAAGGCCCTTGGGAATGGAAGATGAGTTAGTCTGAGTCCGTTTAAATCAG	2608
Db	62	-----	62
QY	2609	AAATCAGGATGAAGGGGTGCGATGACCCGGTTCAAACCTTTTGCACCTGTGGGTCTCG	2668
Db	62	-----	62
QY	2669	GGCCTCACTGCTACCGGCGATGGACCATCATCTGGGAATGGATGCTAACTGGGGCTCT	2728
Db	62	-----	62
QY	2729	CGGCAATTTTGTGACTCTTGCAAGGTTCATACCTGGGTGACGCATCCAAACTGAGTTCT	2788
Db	62	-----	62
QY	2789	CCATCACAAGAGGTGTGACCCCCCTGCCCCACGATCAGAGGCTGGGTCTCTCTCTCT	2848
Db	62	-----	62
QY	2849	TCCACTGCTCACTCTGTTAGCCCGGGGGTGTCTCAAGTTCAAATAGGACTAGGACC	2908
Db	62	-----	62
QY	2909	TCTAGTCTGGGTGATCTCTGGCTTGACAAGAGGCCCTGACCCCTCTGTCAGTTGCGGC	2968
Db	63	-----	66
QY	2969	GCCGCTTCGGGACGTTGTTGACCTGCGCTGGCTGGACCGCGGTGGTCTGCTCAATG	3028
Db	66	ArgTyrGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValLeuLeuAsn	86
QY	3029	GGCTGGCGGCGCTGCGGAGGCGATGTCACCGCGGGGAGGACACGCGCCGCGCCG	3088
Db	86	lyProAlaValIleArgGluAlaLeuValThrTyrGlyGluAspThrAlaAspArgPro	106
QY	3089	CTGCGCCCATCTACCAAGGTCTTGGGCTTGGGCGCGCTTCCCAAGGCAAGCGCGTGG	3148

|||||
106 roAlaHisThrLeuGluProLeuGlyPheGlyProHis----- 118
QY 3149 GGACAGAGACCGCGTTTCCGTGGGCCCCCGGTGGACAGTACCGTAGCCCAAGCAGCGCC 3208
Db 118 ----- 118
QY 3209 GACAGGCGGTGGGCTCTGACGCTGAACAGAGATAAAGCCAGCGAGTGGGCTGAGGAC 3268
Db 118 ----- 118
QY 3269 AGTGGGCCAGAAACACCTGCACGCGGGGAGGTGCGAGTCTGTGGGCTGGGAGGGGGCGG 3328
Db 118 ----- 118
QY 3329 GGCTACTGCCAGACCCCGAGAAGCCCGGTGGGAGGCTGATGCGTGAAGTGGCGGT 3388
Db 118 ----- 118
QY 3389 GCGGGGACCGCGCTATGCTGCGGCTCAGTGTGGCGGGAGCGGGCGGATCTTCCTTG 3448
Db 118 ----- 118
QY 3449 AGTGGAAAGGTGTGAGGTGGGAGAGACGAGTGGGGGCAAAACCCCGCCCGCAGGAGG 3508
Db 118 ----- 118
QY 3509 GGAGCAATGTGGGTGACAAAGAGTGGGCCCTGTGCCACCTGGACCGGCTAGGAGCTG 3568
Db 118 ----- 118
QY 3569 CCGGAGACCTTGTGGAGCGCAGGGTTGGAGTGGGTGGCGAGGGTGGGGCCCAAGGCCCT 3628
Db 118 ----- 118
QY 3629 CATGGCAACGCCACGTGTCCTGCCGCCCGCCAGGGGTGATCTGTGCGCCTATGGGCC 3688
Db 119 ----- 119
QY 3689 GCGTGGCGGAGCAGAGCGCTTCTCGTCTCCACTTGGCACTTGGGCGCTGGGCGCAG 3748
Db 130 AlaTrpArgGluGlnArgPheSerValSerThrLeuArgAsnPheGlyLeuGlyLys 149
QY 3749 AAGTCGTGAGCAGTGGGTGACCGAGGAGCGCGCTTGTGCGGCTTGGCGGACC 3808
Db 150 LysSerLeuGluGlnTrpValThrGluGluAlaThrCysLeuCysAlaAla----- 166
QY 3809 AAGCCGCTGGTGTATGGGCGAAGAGGCGACAAAGCGGAACTGGGAAGCGGGGACGGAG 3868
Db 166 ----- 166
QY 3869 AAGCAACCCCTTACCCGATCTCCCGACCCCGAGCGGCCCTTTCGCCCGCCCAAGCGCT 3928
Db 167 ----- 167
QY 3929 CTTGGCAAAAGCCGTGACCAAGTATGCTCCCTCACCTTGGCACTTGGGCGCTGGGCGA 3988
Db 180 uLeuAsnLysAlaValCysAsnValIleAlaSerLeuThrHisGlyCysArgPheGluTy 200
QY 3989 CGACGACCCCTCGCTTCTCAGGCTGTGGAACCTAGCTCAGGAGGGAAGTGAAGGAGGATC 4048
Db 200 xAspAspHisArgLeuThrArgLeuMetAspLeuThrGlnThrIleLeuLysGluSerTh 220
QY 4049 GGGCTTCTCGCGAGGTGGGAGCGAGACCGAGGAGTCTCTGCGAGGCGGAGCTCCTG 4108
Db 220 rGlyAsnLeuProGln----- 225
QY 4109 AGAGGTCCCGGGCTGACTGGGCGCTCCGAAGGCGCAGGATTTGCATAGATGGTTTGGG 4168
Db 225 ----- 225
QY 4169 AAAGGACATTCAGGAGACCCCACTGTAGAAGGGCTTGGAGGAGGAGGAGGACATCTCAG 4228

Db 225 ----- 225
QY 4229 ACATGTCTGTGGAGAGGTGTGCCCGGTGAGGGGCGCACCGAGAGAGGCCAAGGACTCTG 4288
Db 225 ----- 225
QY 4289 TACCCCGCTCCAGTTGGAGATTTTCGATTTTAGTTTCTCTCTGTGGCAAGGAGAGAGG 4348
Db 225 ----- 225
QY 4349 TGGAGGCTGGCACTTGGGGAGGACTTGGTGAGGTGAGTGGTAAGGACAGGCGAGGCCCTG 4408
Db 225 ----- 225
QY 4409 GGTCTACTGGAGATGGCTGGGCGCTGAGACTTGTCCAGGTGAACGCGAGAGCACAGGAGG 4468
Db 225 ----- 225
QY 4469 GATTGAGACCCCGTTCTGTCTGTGTGAGTGTGAATGTCTCCCGTCTCTCTGCACAT 4528
Db 226 ----- 226
QY 4529 CCCAGCGCTGGTGGCAAGTCTCTACGCTTCCAAAGGCTTCTCTGACCCAGGCTGATGA 4588
Db 236 eProGlyLeuValAspLysValPheArgGlyGlnLysAlaPheMetAlaLeuLeuAspG 256
QY 4589 GCTGTAACTGACGACAGGATGACCTGGGACCCAGCCAGCCAGCCAGGAGCTGACTGA 4648
Db 256 uLeuValThrGluHisArgMetThrArgAspProIleGlnProProArgAspLeuThrAs 276
QY 4649 GGCCTTCTGGCAAGAGGAGAGTGTGAGTGGCTGCCACGCTGGGGGGCAAGGGTGG 4708
Db 276 pAlaPheLeuAspGlnValGluLys----- 284
QY 4709 TGGGTTGAACGTCCTCCAGGAGGAATGAGGGGAGCTGGGCAAAAGGTTGGACAGTGCATC 4768
Db 284 ----- 284
QY 4769 ACCCGCGAGCGCATCTCGGCTGCACAGGTGCAGAAATTGGAGTCTATTTGGGGGTACCC 4828
Db 284 ----- 284
QY 4829 CGTTCTATCCCTGAGTATCTCTCGGCGCTGCTCAGGCCAAGGGAGCCCTGAGAGCAG 4888
Db 285 ----- 285
QY 4889 CTTCAATGATGAACTCGGCATAGTGTGGTGAACCTGTCTCTCGGCGGATGTGAC 4948
Db 292 rPheAsnAspAspAsnLeuValValThrAspLeuPheAlaAlaGlyMetValTh 312
QY 4949 CACTCGACACGCTGGCGCTCTCTGCTCATGATCTACACTGATGTGCAGCG 5008
Db 312 rThrSerIleThrLeuSerTrpAlaLeuLeuMetIleLeuHisProAspValGln-- 331
QY 5009 TGAGCCAGCTGGGGCCCAAGGCGAGGACTGTAGGAGGAGGGGTACAGCTGGGGGGCCCT 5068
Db 331 ----- 331
QY 5069 GGGCTTAGCTGGGACACCCGGGGCTTCCAGACAGGGGTGGCCAGGCTCTGTAAAGCCTA 5128
Db 331 ----- 331
QY 5129 ACTTCTCCACACAGGAGGAGGAGTGTCTCCCTGGGTGCTGACCCATTGTGGGACG 5188
Db 331 ----- 331
QY 5189 CATGTCTGTCCAGTCCGTGTCAAACAGGAGATCGACGACGTGATAGGCGAGGTGCGGCGA 5248
Db 332 ----- 332
QY 5249 CCAGAGATGGGTGACAGGCTCATATGCCCTACACCTGCGCTGATTCACGAGGTGACG 5308
Db 348 ProGluMetGlyAspGlnAlaArgMetProTyThrThrAlaValValHisGluValGln 367

```
QY 5309 CGCTTTGGGACATATCCCTGAGTGTGACCCATATGACATCCGCTGACATCGAAGTA 5368
DB 368 ArgPheAlaAspIleLeuProHisGlnThrSerArgAspIleGluVal 387
QY 5369 CAGGGCTCCGATCCCTAAGGTAGGCTTGGCGCCCTCTCTACCCAGCTCAGCACCGC 5428
DB 388 GlnGlyPheLeuIleProlys----- 394
QY 5429 ACTGGTGTATAGCCCGACGATGGCTACTGCGAGGTGGGGCCACTTAGGAACCTGGCCA 5488
DB 394 ----- 394
QY 5489 CCTAGTCTCAATGCCACACACTGACTGTCCCACTTGGTGGGGGTCCAGAGTATAG 5548
DB 394 ----- 394
QY 5549 GCAGGGCTGGCTGTCCATCCAGAGCCCCGCTAGTGGGGAGACAAACAGGACCTGCC 5608
DB 394 ----- 394
QY 5609 AGAATGTTGAGACCCAGCGCTGCGAGGAGAGGGGCGAGTGTGGTGCCTCTGAGAGG 5668
DB 394 ----- 394
QY 5669 TGTGACTGCGCCTGCTGTGGGTGGAGAGGGTACTGTGGAGCTTCTCGGCGCAGGAC 5728
DB 394 ----- 394
QY 5729 TAGTTGACAGAGTCCAGCTGTGTGCCAGGAGTGTGTCCCGCTGTGTTGGTGGCAG 5788
DB 394 ----- 394
QY 5789 GGGTCCAGCATCTTAGAGTCCAGTCCCACTCTCACCTGCATCTCTGCCAGGGAC 5848
DB 395 -----Gly-Th 396
QY 5849 GACACTCATCCACACCTGCTCGGTGCTGAAGATGAGCGCTGTGGAGAAGCCCTT 5908
DB 396 rValLeuPheThrAsnLeuSerValLeuIysAspGluAlaValTrpGluLysProPh 416
QY 5909 CGCTTTCCACCCCGAACACTTCTGTGATCCCGAGGGCCACTTTGTGAAGCCGAGGCTT 5968
DB 416 eArgPheHisProGlyHisPheLeuAspAlaGlnGlyArgPheValLysGlnGluAlaPh 436
QY 5969 CTGCTCTTTCTCAGCAGGTGCTGTGGGAGCCCGGCTCCCTGTCTCCCTTCCGTGAGTC 6028
DB 436 eMetProPhe--Ser----- 440
QY 6029 TTGCGAGGGTATCACCCAGGACCCAGGCTCACTGACGCCCTTCCCTCCCCACAGCCGC 6088
DB 441 -----AlaGlyArg 443
QY 6089 CGTGCATGCTCGGGAGCCCTGGCCCGCATGGAGCTTCTCTCTTCTTCACTCCCTG 6148
DB 444 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrCysLeu 463
QY 6149 CTGAGCAGCTTCACTTCTCGTGGCCCGGAGACGCCCGCCAGGCCACTCTCTGTGTC 6208
DB 464 LeuGlnArgPheSerPheSerValProThrGlyGlnProArgProSerAspGlnGlyAla 483
QY 6209 GTACGCTTTCTGTGACCCCATCCCTAGAGCTTTGTGCTGTGCCCGC 6259
DB 484 ProAlaThrLeuValThrProAlaProTyrGlnLeuCysAlaValAlaArg 500
```

RESULT 7

```
Q9TUJ5
ID Q9TUJ5 PRELIMINARY; PRT; 500 AA.
AC Q9TUJ5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome P450 2D/1.
```

```
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98391821; PubMed=9722658;
RX Yamamoto Y., Ishizuka M., Takada A., Fujita S.;
RT "Cloning, tissue distribution, and functional expression of two novel
RT rabbit cytochrome P450 isozymes, CYP2D3 and CYP2D24.";
J. Biochem. 124:503-508(1998).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AB008784; BAA84472.1; -.
DR FIR; JE0259; JE0259.
DR HSP; P00179; I076.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; P:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 55721 MW; 7E4A43379F29CEFF3 CRC64;

Alignment Scores:
Pred. No.: 1,01e-68 Length: 500
Score: 1252.50 Matches: 370
Percent Similarity: 29.44% Conservative: 43
Best Local Similarity: 26.37% Mismatches: 84
Query Match: 6.69% Indels: 906
DB: Gaps: 11

US-09-820-788a-3 (1-10278) x Q9TUJ5 (1-500)
QY 2078 ATGGGGCTA-----GAAGCACTGGTGGCCCTGGCCATGATGATGCGCATCTCTCTG 2128
DB 1 MetGlyLeuLeuSerGlyGluAlaLeuAlaProLeuAlaValAlaValAlaPheLeu 20
QY 2129 CTCCTGTGGACCTCATGCACCGCACCAACGCTGGCTGACGCTACCCGCGAGTCCC 2188
DB 21 LeuLeuValAspLeuMetHisLysArgProArgTrpAlaAlaArgLysProGlyPro 40
QY 2189 CTGCCACTGCCCGGCTGGCAACCTTGTGATGTGGACTTCCAGAACACACCATCTG 2248
DB 41 ValGlyIleProGlyLeuGlyAsn-LeuLeuGlnValAspPheArgGlyIleProAsnCy 60
QY 2249 CTTGCACAGGTGAGGAGGAGGTCTCTGGAGGCGGAGAGGTCTGAGGATGCCCCACC 2308
DB 60 sPhe-Arg----- 62
QY 2309 ACCAGCAACATGGTGGTGGTTAAACACAGGCTGGATCAGAGCCAGGCTGAGAGG 2368
DB 62 ----- 62
QY 2369 GGAAGCAGGTTTGGGGGACGTTCTCTGGGGAAGGACATTTATACATGCGATGAAGGACTGG 2428
DB 62 ----- 62
QY 2429 ATTTCCAAAGGCCAAGGAGAGTAGGGAAGGCGCTGGAGGTGGAGCTGGAGCTTGGCAG 2488
DB 62 ----- 62
QY 2489 TGGGCATCAAGCCCATTTGGGCAACATATGTTATGAGGTACAAAGTCCCTCTCTGTGACA 2548
DB 62 ----- 62
QY 2549 CCAGAGGAAGGCGCTTGGGAATGGAAGATGAGTTAGTCTCTGAGTCCGCTTTAAATCAGG 2608
DB 62 ----- 62
```

QY 2609 AAATCGAGGATGAAGGGGGTGCACTGACCCGGTTCAAACCTTTTGGCACTGTGGTCTCTCG 2668
Db 62 ----- 62
QY 2669 GGCCTCACTGTCACCGGCATGGACCATCATCTGGGAATGGATGCTAACTGGGGCTCT 2728
Db 62 ----- 62
QY 2729 CGGCAATTTTGGTGACTCTTGCAAGGTATACCTGGGTGACGCATCCAAACTGATTCCT 2788
Db 62 ----- 62
QY 2789 CCATCAGAAAGTGTACCCCCACCCCTGCCCCACGATCAGGAGGTGGTCTCTCTCT 2848
Db 62 ----- 62
QY 2849 TCCACCTGCTCACTCTGCTAGCCCCGGGGTGTCTCAAGGTTCAATAGGACTAGGACC 2908
Db 62 ----- 62
QY 2909 TGTAGTCTGGGTGATCTGCTTGTCAAGAGGCCCTGACCCCTCTCTCAGTTGCGGC 2968
Db 63 -----GlnLeuArgC 66
QY 2969 GCCGCTTCGGGACGTGTTCAAGCTGAGCTGGCTGGACCGCGTGGTGTCTCAATG 3028
Db 66 ysargTyrGlyAspValPheSerLeuGlnLeuAlaTrpProValValLeuAsnG 86
QY 3029 GGCTGGCGGCTGCGGAGCGATGTGACCGCGCGCGAGGACAGCGCCGACCGCCCGC 3088
Db 86 lyProAlaAlaMetArgGluAlaLeuValThrTyrGlyGluAspThrAlaAspArgProt 106
QY 3089 CTGCGCCCATCTACCAAGTCTCGGCTTCGGCGCGGTCTCCCAAGGCAAGCGCGTGGG 3148
Db 106 yrSerLeuSerLeuGluHisLeuGlyPheGlyProGln----- 118
QY 3149 GGACAGAGACCGGCTTTCCGTGGCGCCCGGTGGACAGTACCGTAGCCCAAGCAGCGCC 3208
Db 118 ----- 118
QY 3209 GACAGGCGTGGGTCTCTGACGTGAACAGAGATAAGGCCAGCGAGTGGGTGAGGAC 3268
Db 118 ----- 118
QY 3269 AGTGGCCAGAAACCACTGCACGGGGAGGTGCGAGTCTGTGGCTGGAGGGGGCGG 3328
Db 118 ----- 118
QY 3329 GGCTACTGCCAGACCGCCAGAAAGCCGGTGGCGAGGCTGATGTCGAAGTGGCGGT 3388
Db 118 ----- 118
QY 3389 GCGGGGACCGCGCTTATGCTGGGGCTCAGTGTGGCGGAGCGGGGGATCTCTCTTG 3448
Db 118 ----- 118
QY 3449 AGTGGAAAGTGTGACGAGTGGCGAGACGAGGTGGGGGCAAAACCCCGCCCGAGGAGG 3508
Db 118 ----- 118
QY 3509 GGAGCAATGTGGTGTAGCAAGAGTGGGCCCTGTGCCACCTGGACCGGGCTAGGAGCTG 3568
Db 118 ----- 118
QY 3569 CGGAGACCTTGTGGAGCGCCAGGTTGGAGTGGGTGGCGAGGGTGGGGCCCAAGGCCCT 3628
Db 118 ----- 118
QY 3629 CATGGCAACGCCCATGCTCCGTCGCCGCCCGAGGGGTGATCTCTGCGCTATGGGCC 3688
Db 119 -----Ala-GlnGlyValIleMetAlaCysTyrGlyHis 129

QY 3689 GCCTGGCGGAGCAGAGGGCGCTTCTCGTGTGCCAATCTTGGCGCTGGGCAAG 3748
Db 130 AlaTrpArgGluGlnArgPheSerValSerThrLeuArgAsnPheGlyMetGlyLys 149
QY 3749 AGTCCGCTGGAGCAGTGGGTGACCGAGGCGCGCTGCTTGTGTCGGCTTCGCCGACC 3808
Db 150 LysSerLeuGluHisTrpValThrGluGluAlaIleCysLeuCysAla-Val----- 166
QY 3809 AAGCCGCTGGTGTATGGGAGAGGCGACAAAGCGGAACTGGGAAAGCGGGGAGCGGAG 3868
Db 166 ----- 166
QY 3869 AAGGCAACCCCTTACCCGCATCTCCCAACCCCGAGCGCCCTTTCGCCCAACCGGCT 3928
Db 167 -----PheSerGluHisAlaGlyHisProPheSerProLysAlaLe 180
QY 3929 CTTGGCAAAACCGTGCAGCAACGTGATCGCTCCCTCCTCAGTGGCGGGCGCGCTTCAGTA 3988
Db 180 uLeuAsnLysAlaIleGlyAsnValIleAlaSerLeuThrPheGlyCysArgPheGluTy 200
QY 3989 CGAGCACCTCGCTTCTCAGGCTGTGACCTAGCTCAGAGGGGACTAAGGAGGAGTC 4048
Db 200 rAspAspHisArgLeuThrArgLeuMetAspLeuIleGluIleMetLeuGluGlnSerTh 220
QY 4049 GGCCTTCTGCGCGAGTCCGGAGCGAGACCGAGGAGTCTCTGCAGGCGGAGCTCTG 4108
Db 220 rGlyIleLeu----- 223
QY 4109 AGAGGTCCGGGCTGGACTGGGGCTCCGAAGGGCAGGATTTGCATAGATGGTGGTGGG 4168
Db 223 ----- 223
QY 4169 AAAGGACATTCAGGAGACCCCACTGTAAAGAGGGCTGGAGGAGGAGGACATCTCAG 4228
Db 223 ----- 223
QY 4229 ACATGTCTGGGAGAGTGTGCCCGGTGAGGGGACCGAGAGGCGCAAGGACTCTG 4288
Db 223 ----- 223
QY 4289 TACCCCGCTCCAGTTGGAGATTTTCATTTTAGTTTCTCTCTGGCAAGGAGAGGG 4348
Db 224 ---ProLeu----- 225
QY 4349 TGGAGCTGGCCTTGGGGGAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4408
Db 225 ----- 225
QY 4409 GGTCTACTGGAGATGGCTGGGGCTGAGACTTGTCCAGGTGAACGAGAGCACAGGAG 4468
Db 225 ----- 225
QY 4469 GATTGAGACCCGTTCTGT 4528
Db 226 -----ValLeuAsnValIleProIleLeuLeuArgI 236
QY 4529 CCCAGCGCTGGTGGCAAGTCTCAGCTTCCAAAGGCTTCTCAGCCCGCTGTGTGTGTGT 4588
Db 236 eProGlyLeuValAspLysValPheHisGlyGlnLysAlaPheMetAlaLeuLeuAspG 256
QY 4589 GCTGCTTAAGTGTGAGCAGGATGACCTGGGACCGAGCCAGCCAGCCAGCCAGCCAGCTG 4648
Db 256 uLeuValThrGluHisArgMetThrArgAspProAlaGlnProProArgAspLeuThrAs 276
QY 4649 GGCCTTCTGGAAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4708
Db 276 pAlaPheLeuAspGlnValGluLys----- 284
QY 4709 TGGGTTGAACGTCGCCAGGAGGATGAGGGGAGCTGGGCAAAAGGTTGGACAGTGCATC 4768
Db 284 ----- 284
QY 4769 ACCCGGAGCGGCATCTGGGCTGACAGGTGACAGATTTGGAGGTTCATTTGGGGGTACCC 4828


```
QY 2111 ATAGTGGCCATCTTCCTGCTCTCTGGTGGAACCTGATGCAACGGGCAACCAACGCTGGGCTGCA 2170
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
15 PheThrValIlePheLeuIleValAsnLeuMetHisTrpArgGlnArgTrpThrAla 34
QY 2171 CGCTACCGCGAGTCCCTGCGCACTGCGCGGCTGGGCAACCTTCTGCTCATGTGACTT 2230
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
35 HisTyProGlyProMetProTrpProValLeuGlyAsn-LeuLeuHisMetAspPh 54
QY 2231 CCAACAACACACCACTACTGCTTCCACCAGGTGAGGAGGAGGTCTCTGGAGGGCGGAGAGG 2290
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
54 eGlnAsnMetPro----- 58
QY 2291 TCCTGAGGATGCCCCCACCACCAAGCAAAACATGGGTGGTGGTTAAACCAACAGGCTGGATCA 2350
Db ----- 58
QY 2351 GAAGCCAGGCTGAGAGGGGAGACAGGTTTGGGGGAGCTTCTGGGGAAGGACATTATATA 2410
Db ----- 58
QY 2411 CATGGCATGAAGACTGGATTTTCCAAAGCCCAAGAGAGTAGGCAAGGGCTGGAGG 2470
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
59 -----AlaGlyPheGln-Lys----- 63
QY 2471 TGGAGCTGACTTGGCAGTGGGCATGCAAGCCCAITGGGCAACATATGTTATGGAGTACA 2530
Db ----- 63
QY 2531 AAGTCCCTTCTGTGACACAGAGAGGAAGGCCCTTGGGAATGGAAGATGATGATGCTCTG 2590
Db ----- 63
QY 2591 AGTGCCGTTTAAATCACGAATCGAGGATGAAGGGGTGCAGTGACCCGCTTCAACCTT 2650
Db ----- 63
QY 2651 TTGCACTGGGTCTCTGGGCTCTCACTGCTCACCGCATGGACCATCATCTGGGAATGG 2710
Db ----- 63
QY 2711 ATGCTAACTGGGCTCTCGGCATTTTGGTGAATTTTGGTGAATCTTGAAGGTATACCTGGGTGAGC 2770
Db ----- 63
QY 2771 CATCAAACCTGAGTTCTCCATCACAGAAAGTGTGACCCCAACCCCTGCGCCACGATCAG 2830
Db ----- 63
QY 2831 GAGGCTGGGTCTCTCTTCCACCTGCTCACTCTGCTAGCCCGGGGCTGCTCCAAGGT 2890
Db ----- 63
QY 2891 TCAATAGGACTAGGACCTGTAGTCTGGGGTGAATCCTGGCTTGACAAGAGGCCCTGACCC 2950
Db ----- 63
QY 2951 TCCCTCTGCAGTTGGCGCCGCTTTCGGGACGTGTTTCAGCTTGACCTGAGCTGGCTGGACGC 3010
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
64 -----LeuArgGlyArgTyArgGlyAspLeuPheSerLeuGlnLeuAlaSerGln 80
QY 3011 CGGTGGTCTGCTCAATGGGCTGGCGCTGCGGAGCGGATGGTGACCCCGCGGAGG 3070
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
80 erValValValLeuLeuGlyLeuThrAlaLeuArgGlnAlaLeuValLysHisSerGln 100
QY 3071 ACAAGCCGACCGCCGCTCGCCCATCTACAGGTCTGGGCTTGGGCGCGCTTCCC 3130
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
100 spThrAlaAspArgProProLeuHisPheAsnAspLeuLeuGlyPheGlyProArgSer- 119
QY 3131 AAGGCAGCGGCTGGGGGACAGACCGCGTTTCCGTGGGCGCGGCTGGACAGTGC 3190
Db :::
120 Gln----- 120
```

```
QY 3191 CGTAGCCCAAGCAGCGCCGACAGGGCGTGGGGTCTTGGAGCGTGAACACAGAGATAAAGGCC 3250
Db ----- 120
QY 3251 AGCGAGTGGGCTGAGGACAGTGGGCCAGGAAACCACCTGCACGGGGGAGGTGCGAGTCTG 3310
Db ----- 120
QY 3311 TGGGCTGGAGGGGGCGGGCTACTGCCACAGACCCGCCAGAAAGCCGCTGGGCGAGGCTG 3370
Db ----- 120
QY 3371 ATGCGTGAAGTGGCGGTGGCGGGACCGCGCCTATGTGTCGGGCTCAGTGTGGGCGGGA 3430
Db ----- 120
QY 3431 CGGGCGGGATCTTCTTGTGAGTGAAGGTGTGAGGTGGGCGAGACGAGGTGGGGCCA 3490
Db ----- 120
QY 3491 AACCCCGCCAGGAGGGGAGCAATGTGGGTGAGCAAGAGTGGGCCCTGTGCCAGCT 3550
Db ----- 120
QY 3551 GGACCGGCTAGGGACTGCGGGAGACCTTGTGGAGCGCCAGGGTTGGAGTGGGTGGCGGA 3610
Db ----- 120
QY 3611 GGGTGGGGCAAGGCCCTTCATGGCAACGCCACGTCGTCCCGCCCCCAGGGGTGATC 3670
Db |||||:::|||||:::|||||:::|||||:::|||||:::
121 -----GlyIleVal 123
QY 3671 CTGTCGCTATGGCGCGCGTGGCGGAGCAGAGCGCTTCTCGTGTCCACCTTGGCGC 3730
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
124 LeuAlaArgTyArgGlyProAlaTrpArgGlnGlnArgPheSerValSerThrMetHis 143
QY 3731 AACTGGGCTGGGCAAGAAGTCGTGGAGCAGTGGGTGACGAGAGGCGCGCTGCCCTT 3790
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
144 HisPheGlyLeuGlyLysSerLeuGlnTrpValThrGluGluAlaArgCysLeu 163
QY 3791 TGTGCGCTTCCCGCACCAAGCCGCTGGTGTGATGGGAGAGGGGCACAAAGCGGAACTG 3850
Db |||||:::|||||:::|||||:::|||||:::|||||:::
164 CysAlaAla----- 166
QY 3851 GGAAGCGGGGACGAGAGGCAACCCCTTACCCGCATCTCCCGACCCCGCAGAGCGCC 3910
Db |||||:::|||||:::|||||:::|||||:::|||||:::
167 -----Phe-AlaAspHisThrGlyHisPr 174
QY 3911 CTTTCGCCCAACCGCTCTTGGACAAAGCCGTGAGCAACGTGATCGCTCCCTCACCTG 3970
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
174 oPheSerProAsnThrLeuLeuAspLysAlaValCysAsnValIleAlaSerLeuLeuTy 194
QY 3971 CGGGCGCGCTTCGAGTACGACACCTCGCTTCTCAGGCTGTGGACTTGTAGCTCAGGA 4030
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
194 rAlaCysArgPheGlyTyAspAspProArgPheIleArgLeuLeuGlyLeuLeuLysG 214
QY 4031 GGAAGTGAAGGAGTGGGCTTCTGCGGAGGTGCGGAGCGAGAGACCGAGGAGTCT 4090
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
214 uThrLeuLysGluGluAlaGlyPheLeu----- 223
QY 4091 CTGACGGGCGAGCTCTGAGAGGTGCCGGGGCTGACTGGGGCTCCGAAAGGCGAGATT 4150
Db ----- 223
QY 4151 TGCATAGATGGTTTGGAAAGGACATTCAGAGAGACCCCACTGTAAGAAGGCGCTGGAG 4210
Db ----- 223
QY 4211 GAGAGGGGACATCTCAGACATGTCGTGGGAGAGGTGTGCCCGGCTCAGGGGGCACCCAG 4270
Db ----- 223
QY 4271 GAGAGGCAAGGACTCTGTATACCCCGTCCACGCTTGAGATTTCGAGATTTTAGTTTCTCCT 4330
```

```
Db 224 -----ProMet----- 225
QY 4331 CTGGCAAGAGAGAGAGGCTGGAGGCTGGCACTTGGGGAGGGACTTGGTGAGGTCAAGTGGT 4390
Db 225 ----- 225
QY 4391 AAGGACAGGAGGCGCTGGTCTACCTGGAGATGGCTGGGGCTGAGACTTGTCCAGGTG 4450
Db 225 ----- 225
QY 4451 AACGCAGAGCACAGGAGGATTTAGACACCCGTTCTGTCTGTGTAGGTCTGAATGCTGT 4510
Db 226 -----PheLeuAsnValPh 230
QY 4511 CCCCTCTCTCCGACATCCACAGCCCTGGCTGGCAAGTCTTACGCTTCCAAAGGCTTT 4570
Db 230 eProMetLeuLeuAArgGleProGlyLeuValGlyLysValPheProGlyLysArgAlaPh 250
QY 4571 CCTGACCCAGCTGAGTGAAGTCTAACTGACGACAGGATGACCTGGGACCCAGCCAGCC 4630
Db 250 eValThrMetLeuAspGluLeuLeuAlaGluHisLysThrThrTrpAspProThrGlnPr 270
QY 4631 ACCCCGAGACCTGACTGAGGCTTCTCTGGCAAGAGAGAGAGTGGAGTGGCTGCCAC 4690
Db 270 oProArgAspLeuThrAspAlaPheLeuAlaGluValGluLys----- 284
QY 4691 GGTGGGGGCAAGGTGTGGTGTGAACGTCCAGGAGGAATGAGGGAGGCTGGGCAAA 4750
Db 284 ----- 284
QY 4751 AGTTGGACCATGTCATCACCGGCGAGCGCATCTGGGCTGACAGGTGCAGATTGGAG 4810
Db 284 ----- 284
QY 4811 GTCAATTGGGGCTACCCGTTCTATCCCTGAGTATCTCTCGCCCTGCTCAGGCCAA 4870
Db 285 -----AlaIy 286
QY 4871 GGGGAGCCCTGAGACAGCTTCAATGATGAGAACCTCGCGCATAGTGGTGGTAACTCTTT 4930
Db 286 eGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgThrValValGlyAspLeuPh 306
QY 4931 CCTTCCCGGATGTGACACTCGACACAGCTGGCTGGGCTCTCTGCTCATGATCCT 4990
Db 306 eSerAlaGlyMetValThrSerThrLeuSerThrLeuSerThrPalaLeuMetLeuMetIleLe 326
QY 4991 ACACCTGGATGTGACCGCTGAGCCAGCTGGGGCCCAAGCGAGGAGTGTGAGGAGGAAG 5050
Db 326 uHisProAspValGln----- 331
QY 5051 GTACAGTGGGGGCCCCCTGGGCTTAGCTGGGACACCCCGGGCTTCCAGCACAGGCGTGGC 5110
Db 331 ----- 331
QY 5111 CAGGCTCTGTAAAGCTAACTTCTCCACACAGGAGGAAGAGTGTCCCTGGGTGC 5170
Db 331 ----- 331
QY 5171 TGACCAATTGTGGGACGATGTCGTCCAGTCCGTGTCACAGAGAGATCGACGAGCTG 5230
Db 332 -----Arg-ArgValGlnGlnGluLeuAspGluVal 341
QY 5231 ATAGGCGAGGTGCGGCGACAGAGATGGGTGACAGGCTCATGCTCCCTACACCACTGCC 5290
Db 342 IleGlyGlnValGlnCysProGluMetAlaAspGlnAlaArgMetProFyrThrAsnAla 361
QY 5291 GTGATTCAGAGGTGACGGCTTTGGGACATCATCCCTGAGTGTGACCCATATGACA 5350
Db 362 ValIleHisGluValGlnArgPheAlaAspIleLeuProLeuGlyValProHisLysThr 381
QY 5351 TCCCGTGCATCGAAGTACAGGCTTCCGCTATCCCTAAGGTAGGCTGGCGCTCTCTCA 5410
|||||
```

```
382 SerArgAspIleGluLeuGlnGlyPheLeuIleProLys----- 394
QY 5411 CCCAGCTCAGCACACGACACCTGGTGTAGTATAGCCCGAGCATGGCTACTGCCAGTGGGCCCA 5470
Db 394 ----- 394
QY 5471 CTCTAGGAACCTGGCCACCTAGTCTCAATGCCACACACACTGACTGTCCCACTTGGT 5530
Db 394 ----- 394
QY 5531 GGGGGTCCAGAGTATAGGAGGGCTGGCTGTCTCATCCAGAGCCCCCGTCTAGTGGGA 5590
Db 394 ----- 394
QY 5591 GACAAAACAGGACCTGCCAGAAATGTTGGAGACCAGCCGCTGCAGGAGAGGGGGCAGT 5650
Db 394 ----- 394
QY 5651 GTGGGTGCTCTGAGAGGTGTGACTGGCCCTGCTGTGGGGTGGGAGAGGTACTGTGGA 5710
Db 394 ----- 394
QY 5711 GCTTCTCGGCGCAGGACTAGTTGACAGTCCAGCTGTGTGCCAGGAGTGTGTGCC 5770
Db 394 ----- 394
QY 5771 CCGTGTGTTGGTGGCAGGGGTCCAGCATCTAGATCCAGTCCCACTCTCACCTGC 5830
Db 394 ----- 394
QY 5831 ATCTCTGCCCCAGGAAACGACACTCATCACCACCTGTCTCATCGTGTGAAGGATGAGC 5890
Db 395 -----Gly- ThrThrLeuIleThrAsnLeuSerSerAlaLeuLysAspGluTh 410
QY 5891 CGTCTGGAGAGCCCTTCCGCTTCCACCCCGAACACTTCTCTGGATGCCAGGCGCACTT 5950
Db 410 rValTrpGluLysProLeuCysPheHisProGluHisPheLeuAspAlaGlnGlyHisPh 430
QY 5951 TGTGAGCGGAGAGCTTCTCTGCTTCTCAGCAGGTGCTGTGGGGAGCCGGCTCCT 6010
Db 430 eValLysProGluAlaPheMetProPhe--Ser----- 440
QY 6011 GTCCCTTCCGTGGAGTCTTGCAGGGGTATCACCCAGGAGCAGGCTCACTGACGCCCT 6070
Db 440 ----- 440
QY 6071 CCCCTCCACAGCCCGCTGCATGCTCGGGAGCCCTGGCCCGCATGAGCTCTTC 6130
Db 441 -----AlaGlyArgArgSerCysLeuGlyGluProLeuAlaArgMetGluLeuPhe 457
QY 6131 CTCTTCTTCACTCCCTGCTGCTGAGCTTCTGAGCTTCTCCGTGGCGCGGACAGCCCGG 6190
Db 458 LeuPhePheThrCysLeuLeuGlnArgPheSerIleSerValProAspGlyGlnProGln 477
QY 6191 CCCAGCCACTCTCGTGTGCTGAGCTTCTGTGTGACCCCATCCCTCAGAGCTTGTGCT 6250
Db 478 ProSerAspHisGlyValPheArgAlaLeuThrThrProCysProFyrGlnLeuCysAla 497
QY 6251 GTGCCCGGC 6259
Db 498 LeuProArg 500
RESULT 9
QY1W87 PRELIMINARY; PRT; 500 AA.
AC QY1W87
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
GN Cytochrome P450, 2d22.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```


QY 3791 TGTGCGCTTCCTCCGACCAAGCCGTTGGTGTATGGGCAGAGGGCACAAAGCGGAACGTG 3850
Db |||||
164 CysAlaAla----- 166
QY 3851 GGAAGCGGGGACGAGAGCAACCCCTTACCCTGATCTCCACACCCCGAGAGCGCC 3910
Db |||||
167 -----Phe-AlaAspHisThrGlyHisPr 174
QY 3911 CTTTCGCCCAAGCGCTCTTCGACAAAGCCGTGAGCAACGTGATCGCTCCCTCACCTG 3970
Db |||||
174 oPheSerProAsnThrLeuLeuAspLysAlaValCysAsnValIleAlaSerLeuLeuTy 194
QY 3971 CGGGCGCGCTTCGAGTACGACACCTCGCTTCCTCAGGCTGCTGGACCTAGCTCAGGA 4030
Db |||||
194 rAlaCysArgPheGluTyAspAspProArgPheIleArgLeuLeuGlyLeuLeuLysG 214
QY 4031 GGGACTGAAGAGAGTGGGCTTTCTGCGGAGTGGCGAGCGGAGACCGAGGAGTCT 4090
Db |||||
214 uThrLeuLysGluGluAlaGlyPheLeu----- 223
QY 4091 CTGACGGCGAGCTCTGAGAGGTGCGGGGCTGACTGGGGCTCCGAGGGCAGGATT 4150
Db ----- 223
QY 4151 TGCATAGATGGTTTGGGAAAGGACATTCAGAGACCCCACTGTAAAGAGGGCTGGAG 4210
Db ----- 223
QY 4211 GAGGAGGGACATCTCAGACATGTTGTTGGGAGAGTGTGCCGGTTCAGGGGGCACGAG 4270
Db ----- 223
QY 4271 GAGAGGCCAAGGACTCTGTACCCCGTCCAGTTGGAGATTTTCGATTTTAGTTTCTCT 4330
Db |||||
224 -----ProMet----- 225
QY 4331 CTGGCAAGAGAGAGGGTGGAGGCTGGCACTTGGGGAGGACTTGGTGAGTCAAGTGT 4390
Db ----- 225
QY 4391 AAGGACGGAGCGCCCTGGTCTACCTGGAGATGGCTGGGGCTGAGACTTGTCCAGGTG 4450
Db ----- 225
QY 4451 AACGACAGACAGAGGAGATTGAGACCCGTTCTGTCTGTGTAGTGTCTGAATGTGT 4510
Db |||||
226 -----PheLeuAsnValPh 230
QY 4511 CCCGCTCTCTGCACATCCAGCGCTGGCTGGCAAGGTCTACGCTTCCAAAAGGCTTT 4570
Db |||||
230 eProMetLeuLeuArgIleProGlyLeuValGlyLysValPheProGlyLysArgAlaPh 250
QY 4571 CTTGACCCAGCTGGATGACTCTAACTGAGCACAGGATGACTGGGACCCAGCCAGCC 4630
Db |||||
250 eValThrMetLeuAspGluLeuLeuAlaGluHisLysThrTrpAspProThrGlnPr 270
QY 4631 ACCCGACACCTGACTGAGGCTTCTGCGAAAGAGAGAGAGTGGAGTGGTGGCCAC 4690
Db |||||
270 oProArgAspLeuThrAspAlaPheLeuAlaGluValGlyLys----- 284
QY 4691 GGTGGGGGCAAGGTGTGGTGTGAACGTCCCAGGAGGAATGAGGGGAGGCTGGGCAAA 4750
Db ----- 284
QY 4751 AGTTTGGACAGTGCATCACCGGCGAGCCGATCTGGGCTGACAGGTGCAGAAATTGGAG 4810
Db ----- 284
QY 4811 GTCATTTGGGGCTACCCGTTCTATCCCTGAGTATCCTCTGGGCCCTGTCTAGGCCAA 4870
Db |||||
285 -----AlaLy 286
QY 4871 GGGAGCCCTGAGAGCAGCTTCAATGATGAGAACCTGGCATAGTGTGGGTAACTGT 4930

Db sGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgThrValValGlyAspLeuPh 306
QY 4931 CTTTCGCCGATGTGACCACTGCACACCGCTGGCGCTTCCTGCTCATGATCCT 4990
Db |||||
306 eSerAlaGlyMetValThrThrSerThrLeuSerTrpAlaLeuMetLeuMetIleLe 326
QY 4991 ACACCTGATCTGCAGCGTGTAGCCAGCTGGGGCCCAAGGAGGAGCTGAGGGAGGAGG 5050
Db |||||
326 uTyProaspValGln----- 331
QY 5051 GTACAGCTGGGGGCGCTTGGGCTTAGTGGGACACCCGGGCTTCCAGCACAGCGGTGGC 5110
Db ----- 331
QY 5111 CAGGCTCCTGTAAAGCTTAACCTTCTCCAAACACAGGAGGAGAGTGTCCCTGGGTGC 5170
Db ----- 331
QY 5171 TGACCCATTGTGGGACGCGATGTCTGTCCAGTCCGTCTCAACAGGAGATCGACGACGTG 5230
Db |||||
332 -----Arg-ArgValGlnGlnGluIleAspGluVal 341
QY 5231 ATAGGGCAGGTGCGCGACACAGAGATGGGTGACAGGCTCACATGCCCTCACCACTGCC 5290
Db |||||
342 IleGlyGlnValGlnCysProGluMetAlaAspGlnAlaArgMetProTyThrAsnAla 361
QY 5291 GTGATTACGAGGTGACGCGCTTTGGGGACATCATCCCTGAGTGTGACCATATGACA 5350
Db |||||
362 ValIleHisGluValGlnArgPheAlaAspIleLeuProLeuGlyValProHisLysThr 381
QY 5351 TCCGCTGACATCGAAGTACAGGCTTCGGCATCTCCTAAGGTAGGCTGGCGCCCTCCTCA 5410
Db |||||
382 SerArgAspIleGluLeuGlnGlyPheLeuIleProLys----- 394
QY 5411 CCCAGCTCAGACACCGACCTGGTGTAGTCCGCCAGCATGGCTACTGCGAGTGGGCCCA 5470
Db ----- 394
QY 5471 CTGTAGAACCTCGGCCACTAGTCTCTCAATGCCACCACTGACTGTCTCCACTTGGGT 5530
Db ----- 394
QY 5531 GGGGGTCCAGATATAGGACAGGCTGGCTGTCTCCATCCAGAGCCCCGCTCTAGTGGGA 5590
Db ----- 394
QY 5591 GACAAACAGGACCTGCCAGAAATGTTGGAGGACCCAGCGCTGCAGGAGAGGGGAGT 5650
Db ----- 394
QY 5651 GTGGTGCCTCTGAGAGGTGTGACTGCGCCCTGCTGTGGGTGGAGAGGAGTCTGTGGA 5710
Db ----- 394
QY 5711 GCTTCTCGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAGGAGTGTGTGTC 5770
Db ----- 394
QY 5771 CCGTGTGTTTGGTGGCAGGGTCCAGCATCTTAGAGTCCAGTCCCTCACCCTGC 5830
Db ----- 394
QY 5831 ATCTCTGCCCGGAAACGACACTCATCAACACTCTCATCGTGTGTGAGGATGAGGC 5890
Db |||||
395 -----Gly-ThrThrLeuIleThrAsnLeuSerSerAlaLeuLysAspGluTh 410
QY 5891 CGTCTGGAGAGCCCTTCGCTTCCACCCCGAAACACTTCTCTGTGATGATGCCAGGCCACTT 5950
Db |||||
410 rValTrpGluLysProLeuCysPheHisProGluHisPheLeuAspAlaGlnGlyHisPh 430
QY 5951 TGTGAACCGGAGGCTTCTGCTTCTCAGCAGGTGCTGTGTGGGAGCGCCGCTCCT 6010
Db |||||

```

Db      430 eVallysProGluAlaPheMetProPhe--Ser----- 440
QY      6011 GTCCCTTCCGTGGAGTCTTGAGGGGTATCACCCAGGAGCCAGGCTCACTGACGCCCT 6070
Db      440 ----- 440
QY      6071 CCCCTCCACAGGCGCGTGCATGCTCGGGGAGCCCTGGCGCGGATGGAGCTCTTC 6130
Db      441 -----AlaGlyArgArgSerCysLeuGlyGluProLeuAlaArgMetGluLeuPhe 457
QY      6131 CTCTTCTTCACTCCCTGCTGAGCAGCTTTCAGCTTCTCGTGGCGCGCGGAGCCCGG 6190
Db      458 LeuPhePheThrCysLeuLeuGlnArgPheSerIleSerValProAspGlyGlnProGln 477
QY      6191 CCAGCCACTCTCTGTCGTGCTGAGCTTCTGTCGACCCCATCCCTACGAGCTTTGTGCT 6250
Db      478 ProSerAspHisGlyValPheArgAlaLeuThrThrProCysProTyGlnLeuCysAla 497
QY      6251 GTGCCCGCGC 6259
Db      498 LeuProArg 500

RESULT 10
Q8VCX0 PRELIMINARY; PRT; 500 AA.
AC Q8VCX0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to cytochrome P450, 2d9.
GN 1300007K12RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; BC018344; AAH18344.1; -.
DR MGD; MGI:1915694; 1300007K12RIK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR001128; Cytochrome P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00867; CPASE 2; 1.
DR PROSITE; PS00866; CYTOCHROME P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 56625 MW; 30017F7A1AE4PB3E CRC64;

```

Alignment Scores:

```

Pred. No.: 9,14e-64 Length: 500
Score: 1172.00 Matches: 344
Percent Similarity: 28.82% Conservative: 58
Best Local Similarity: 24.66% Mismatches: 90
Query Match: 6.26% Indels: 903
DB: 11 Gaps: 10

```

US-09-820-788a-3 (1-10278) x Q8VCX0 (1-500)

```

QY      2093 CTGGTGCCCTGGCCATGATAGTGGCCATCTTCTGCTCTGTCGACCTGATGACCGG 2152
Db      9 LeuTrpProValAlaIlePheThrValIlePheIleLeuLeuValAspLeuMethHisArg 28
QY      2153 CACCAACGCTGGGCTGACAGCTACCGCGCAGGCTCCCTGCCACTGCGCGGCTGGGCAAC 2212

```

```

Db      29 ArgGlnArgTrpThrSerArgTyrProGlyProValProTyrProValLeuGlyAsn 48
QY      2213 CTTGTGTCATGTGGACTTCCAGACACACACCACTACTGCTTCGAGCTGAGGTGAGGAGGT 2272
Db      49 -LeuLeuGlnValAspLeuAspAsnMetProTyr-Ser----- 60
QY      2273 CTTGAGGCGCGCAGAGGTCTTGAGGATGCCCCACACAGCAAAATGCGGTGGTGGT 2332
Db      60 ----- 60
QY      2333 AAACACAGGCTGATCAGAACCCAGGCTGAGAAGGGAGCAGGTTTGGGGGACGTTCC 2392
Db      60 ----- 60
QY      2393 TGGGAAGGACATTTATATATGTCATGAGGACTGGAATTTCCAAAGGCCAAGGAAGT 2452
Db      61 -----LeuTyr----- 62
QY      2453 AGGCAAGGCGCTGGAGGTGGAGCTGGACTTGGCAGTGGGCATGCAAGCCCATTTGGCAA 2512
Db      62 ----- 62
QY      2513 CATATGTTATGAGGTACAAAGTCCCTTCTGTGACACCAAGGAAGGCGCTTGGGATG 2572
Db      62 ----- 62
QY      2573 GAAGATGAGTTAGTCTCTGAGTCCGTTAAATACGAAATCGAGGATGAAGGGGTGCAG 2632
Db      62 ----- 62
QY      2633 TGACCCGGTTCAAACCTTTTGACACTGTGGGTCTCTCGGCGCTCACTGCTCACCGGCATGA 2692
Db      62 ----- 62
QY      2693 CCATCATCTGGGATGGGATGCTAACTGGGGCTCTCGGCAATTTTGGTGACTCTTGCAA 2752
Db      62 ----- 62
QY      2753 GGTATACCTGGTGACGCATCCAAACTGAGTTCTCCATCACAGAAGGTGTGACCCCA 2812
Db      62 ----- 62
QY      2813 CCCCTGCCACGATCAGGAGGCTGGGTCTCTCTTCCACCTGCTCACTCTGCTAGCC 2872
Db      62 ----- 62
QY      2873 CCGGGGCTGTCACAGGTTCAAAATAGGACTAGGACCTGTAGTCTGGGGTGTATCTGGCTT 2932
Db      62 ----- 62
QY      2933 GACAAGAGCCCTGACCCCTCTCTGAGTTGCGGCGCGCTTCCGGGAGCGTTTCAGCC 2992
Db      63 -----LysLeuGlnAsnArgTyrGlyAspValPheSerL 74
QY      2993 TGCAGCTGGCCTGGACCGCGGTGGTGGTGGTCAATGGGTGGCGCGCTGCGCGAGGCCA 3052
Db      74 euGlnMetAlaTrpLysProValValIleSerGlyLeuLysAlaValArgGluValL 94
QY      3053 TGGTGACCCGCGGCGAGGACACGCGCCGACCGCGCTGCGCCCTATCTACAGGCTCTGG 3112
Db      94 euValThr-CysGlyGluAspThrAlaAspArgProGluMetProIlePheGlnHisLeuG 114
QY      3113 GCTTCGGGCGCGTTCCTCCAAAGCAAGCGCGGTGGGGGACAGACCGCGTTTCCGTGGG 3172
Db      114 lyTyrglyglulysAla-Lys----- 120
QY      3173 CCGCGGCTGGACAGTGAAGCCGCTAGCCCAAGCAGCGCGCGAGGCGCTGGGGTCTTGAC 3232
Db      120 ----- 120
QY      3233 GAAACAGAGATAAGGCCAGCGAGTGGGCTGAGGACAGTGGGCCAGGAAACCACTGAC 3292

```

Db 120 ----- 120
Qy 3293 GGGGAGGTGCGAGTCTGTGGCTGGGAGGGGGGGGCTACTGCCAGACCCGCGAGAA 3352
Db 120 ----- 120
Qy 3353 GCCCGTGGCGAGGCTGATCGTTCGAGTGGCGGTGGCGGGACCGCGCTATGCTGG 3412
Db 120 ----- 120
Qy 3413 GGCTCAGTGTGGCGGAGCGCGGGATCTTCTTTAGTGGAAAGGTGGTCAGGGTGGGC 3472
Db 120 ----- 120
Qy 3473 AGAGACGAGGTGGGCCAAACCCCGCCAGCAGGGAGCAATGTGGTGAGCAAGAG 3532
Db 120 ----- 120
Qy 3533 TGGGCCCTGTGCCAGCTGGACCGGGCTAGGAGCTGCGGAGACCTTGTGGAGCCAGG 3592
Db 120 ----- 120
Qy 3593 GTTGAGTGGTGGCGGAGGGTGGGCCAAGGCCCTTCATGGCAACGCCCAAGTGTCCGTC 3652
Db 120 ----- 120
Qy 3653 CCGCCCCAGGGGTGATCTCTGCGCTATGGGCCCGCGTGGCGCGAGCAGGGCGCTTC 3712
Db 121 ----- GlyValValPheAlaProTyrGlyProGluTyrArgGluLeuArgPhe 137
Qy 3713 TCCGTGTCCACCTTCGCAACTTGGCTGGCAAGAGTGGCTGGAGCAGTGGGTGACC 3772
Db 138 SerValSerThrLeuArgAsnLeuGlyLeuGlyLysSerLeuGluGlnTrpValThr 157
Qy 3773 GAGGAGCGCGCTGCTTTGTGCCCTTCGCGCACCAAGCCGTTGATGGGCGAGAAG 3832
Db 158 GluGluAlaGlyHisLeuCys ----- Asp-AlaPheTh 168
Qy 3833 GCACAAAGCGGAATCGGAAGCGGGGACGAGAAAGCAACCCCTTACCCGATCTC 3892
Db 168 rAlaGlnAlaGlySer ----- 173
Qy 3893 CCCACCCAGGACGCCCTTTGGCCCAACGGCTCTTGGACAAAGCGGTGAGCAAGT 3952
Db 174 ----- ProLeuAspProTyrThrLeuLeuAsnLysAlaValCysAsnVa 188
Qy 3953 GATCCCTCCCTCACCTGGCGGCGCGCTTCGAGTACGACGACCTCGCTTCTCAGGCT 4012
Db 188 lIleAlaSerLeuIleTyrAlaArgArgPheGluTyrGlyAspProAspPheIleLysMe 208
Qy 4013 GCTGACCTAGCTCAGGAGGACTGAAGGAGGAGTCGGGCTTTCTGCGGAGGTGCGGAG 4072
Db 208 tLeuLysIleLeuLysGluAsnMetGlyGluAsnThrGlyLeuPheProGlu- 225
Qy 4073 CGAGACCGGAGGTCTCTGCGGCGGAGCTCCTGAGAGGTGCGCGGCTGAGTGGG 4132
Db 225 ----- 225
Qy 4133 CCTCCGAAGGCGAGGATTTCATAGATGGTTTGGGAAAGGACATTCAGGAGACCCAC 4192
Db 225 ----- 225
Qy 4193 TGTAAAGGCGCTGGAGGAGGGGACATCTCAGACATGTTGTGGGAGAGGTGTGCC 4252
Db 225 ----- 225
Qy 4253 CGGGTCAGGGGACACAGGAGGCCAAGGACTCTGTATCCCGCTCCAGTTGGAGATT 4312
Db 225 ----- 225
Qy 4313 CGATTTTAGTTTCTCTCTGGGCAAGGAGAGGGTGGAGGCTGGCACTTGGGAGGGA 4372
Db 225 ----- 225

Qy 4373 CTTGGTGAAGTTCAGTGGTAAGCAGGACAGCCCTTGGGTCTACCTGGAGATGGCTGGGC 4432
Db 225 ----- 225
Qy 4433 CTGAGACTTGTCCAGGTGAACGACAGACAGGAGGATTGAGACCCCGTTCTGTCTGGT 4492
Db 225 ----- 225
Qy 4493 GTAGTGTGTAATGTCTCCCGTCTCTCTGCACATCCAGCGTGGCTGGCAAGGTCT 4552
Db 226 ---ValLeuAsnThrPheProIleLeuLeuHisIleProGlyLeuAlaAspLysValph 244
Qy 4553 ACGCTTCCAAAAGGTTCTCTGACCCAGCTGGATGAGTGTCTAACTGACGACGATGAC 4612
Db 244 eProGlyGlnLysThrPheLeuThrLeuValAsnLysLeuValThrGluHisLysArgTh 264
Qy 4613 CTGGACCCAGCCAGCACCCAGACCTGACTGAGGCTTCTCTGGCAAGAGAGAGAA 4672
Db 264 rTTPAspProaspGlnProProArgaspLeuThrAspAlaPheLeuAlaGluMetGluLys 284
Qy 4673 GGTGAGAGTGTGCTGCCACGTTGGGGGCAAGGGTGGTGGTGAACGTCCACGAGGAAT 4732
Db 284 s----- 284
Qy 4733 GAGGGAGGCTGGGCAAAAGGTTGGACCAAGTGCATACCCGGCGAGCGCATCTGGGCTG 4792
Db 284 ----- 284
Qy 4793 ACAGGTGCAGAAATTGGAGTCAATTGGGGGCTACCCCGTTCTATCTCCCTGAGTATCTCT 4852
Db 284 ----- 284
Qy 4853 CGGCCCTGCTCAGGCCAAAGGGAGCCCTGAGAGCAGCTTCAATGATGAGAACCTGCGCAT 4912
Db 285 ---AlaLysGlyAsnProLysSerSerPheAsnGluAlaAsnLeuArgLe 300
Qy 4913 AGTGGTGGTAACCTGTCTCTGGCGGATGGTGCACCTCGACCCAGCTGGCTGGG 4972
Db 300 uValValPheAspLeuPheGlyAlaGlyIleValThrSerSerIleThrLeuThrTrpAl 320
Qy 4973 CCTCTGCTCATGATCTACACCTGGATGTGACGCTGAGCCAGCTGGGGCCCAAGGCA 5032
Db 320 aLeuLeuLeuMetIleLeuHisProaspValGln----- 331
Qy 5033 GGGACTGAGGAGGAAGGTTACAGTGGGGGCCCTTGGGCTTAGCTGGGACACCCGGGGC 5092
Db 331 ----- 331
Qy 5093 TTCCAGCACAGCGGTGGCCAGGCTCTGTAAAGCTTAACCTCTCCACACAGGAGGAAG 5152
Db 331 ----- 331
Qy 5153 AGAGTGTCCCTGGTGTGTGACCCATTGTGGGAGCGATGTCTGTCCAGTCCGTGTCAA 5212
Db 332 ---Arg-ArgValGln 335
Qy 5213 CAGAGATTCAGCAGATGATAGGCGAGGTGGCGCACAGAGATGGGTGACAGGCTCAC 5272
Db 336 GluGluIleaspGluValIleGlyGlnValArgCysProGluMetAlaAspGlnAlaHis 355
Qy 5273 ATGCCCTACACCTGCGGTGATTCAGAGGTGCGGCTTGGGGACATCATCCCTG 5332
Db 356 MetProTyrThrAsnAlaValIleHisGluValGlnArgPheAlaAspIleValProMet 375
Qy 5333 AGTGTGACCATATGACATCCCGTGCATCAGAGTACAGGCTTCCGAGTCCCTAAGTA 5392
Db 376 AsnLeuProHisLysThrSerHisAspIleGluValGlnGlyPheLeuIleProLys--- 394
Qy 5393 GGCCTGGCGCCCTCTCCACCCAGCTCAGCACAGCACCTGGTGTATAGCCGCCAGCATGGC 5452
Db 394 ----- 394

QY 5453 TACTGCCAGGTGGGCCCCACTCTAGGAACCTGTGCCACCTAGTCTCTCAATGCCACCACT 5512
DB 394 ----- 394
QY 5513 GACTGTCCCCACTTGGTGGGGGTCCAGAGTATAGCAGGGTGCCTGTCCATCCAGA 5572
DB 394 ----- 394
QY 5573 GCCCCGCTCTAGTGGGGAGACAAACCAGGACCTGCCAGAAATGTGTGGAGACCAGCGCT 5632
DB 394 ----- 394
QY 5633 GCAGGAGAGGGGGCAGTGTGGTGTCTGTGAGAGGTGTGACTGCCCTGTGTGGGT 5692
DB 394 ----- 394
QY 5693 CGGAGAGGTACTGTGGAGCTTCTCGGGCAGGACTAGTTGACAGAGTCCAGCTGTGTG 5752
DB 394 ----- 394
QY 5753 CCAGGCAGTGTGTCTCCCGTGTGTGTGTGGTGGCAGGGTCCAGCATCTAGAGTCCAG 5812
DB 394 ----- 394
QY 5813 TCCCCACTCTCACCTGTGCATCTCTGCCAGGAAACGACACTCATCACCACTGTGTATC 5872
DB 395 ----- Gly-ThrThrLeuLeuProAsnLeuSerSe 404
QY 5873 GGTGTCTGAGGATGAGCGCTCTGGGAGAGCCCTTCGGTTCACCCCGGACACTTCCT 5932
DB 404 rThrLeuLysAspGluThrValTrpGluLysProLeuArgPheHisProGluHisPheLe 424
QY 5933 GATGTCGAGGGGCACTTGTGTGAAGCGGAGGCTTCTCGCTTCTCTGAGCAGGTGCCTG 5992
DB 424 uAspAlaGlnGlyHisPheValLysProGluAlaPheMetProPhe--Ser----- 440
QY 5993 TGGGGAGCGCGGCTCCCTGTCTCCCTTCCGTGGAGTCTTGCAGGGGTATCACCAGAGCC 6052
DB 440 ----- 440
QY 6053 AGGCTCACTGACGCCCTCCCTCCCAAGCGCGCGTGCATGCTCGGGAGCCCTG 6112
DB 441 ----- AlaGlyArgAlaCysLeuGlyProLeu 451
QY 6113 GCCCGCATGAGCTTCTCTCTTCTTCTCACCTCCCTGCTGAGCAGCTTCTGCTCCGTG 6172
DB 452 AlaArgMetGluLeuPheLeuPheThrCysLeuLeuGlnArgPheSerPheLeuVal 471
QY 6173 GCGCGGAGAGCCCCGCGCCAGCCACTCTCGTGTCTGCTGAGCTTCTGTGAGCCCATCC 6232
DB 472 ProAlaGlyGlnProGlnProSerAspTyrGlyIlePheThrPheLeuValSerProSer 491
QY 6233 CCTACGAGCTTGTGTGTGCGCCCGC 6259
DB 492 ProTyrGlnLeuCysAlaPheThrArg 500

RESULT 11

Q8CIM7 PRELIMINARY; PRT; 500 AA.
AC Q8CIM7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RIKEN cDNA 1300006E06 gene.
GN CYP2D6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RA Strausberg R.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC023241; AAH23241.1; -.
DR MGI; MGI:1923529; Cyp2d26.
DR GO; GO:0016712; F-oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P-electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
SQ SEQUENCE 500 AA, 56975 MW, B3DDCD88DFA3F265 CRC64;
Alignment Scores:
Pred. No.: 2,3e-63 Length: 500
Score: 1165.50 Matches: 348
Percent Similarity: 28.80% Conservative: 56
Best Local Similarity: 24.80% Mismatches: 93
Query Match: 6.22% Indels: 906
DB: 11 Gaps: 10
US-09-820-788A-3 (1-10278) x Q8CIM7 (1-500)

QY 2078 ATGGGGCTA-----GAAGCACTGGTGGCCCTGGCCATGATAGTGGCCATCTCTCTG 2128
DB 1 MetGlyLeuLeuValGlyAspLeuTrpAlaValValIlePheThrAlaIlePheLeu 20
QY 2129 CTCTGTGTGACCTGATGCACCGCACCAACCTGGCTGGCTGACGCTTACCAGGAGTCCC 2188
DB 21 LeuLeuValAspLeuValHisArgGlnArgGlnArgTrpThrAlaCysTrpProGlyPro 40
QY 2189 CTGCCACTGCCCGGCTGGGCAACCTTGTGTGATGTGACTTCCAGAAACACCATCTG 2248
DB 41 ValProPheProGlyLeuGlyAsn-LeuLeuGlnValAspPheGluAsnIleProTyrSe 60
QY 2249 CTTGACACAGTGTGAGGAGGAGGTCTCTGGAGGGCGGACAGAGTCTGTGAGGATGCCCCACC 2308
DB 60 rPhe-TyrIys----- 63
QY 2309 ACCAGCAACATCGGTGGTGGTTAAACACACAGGTGGATCAGAACCCAGGCTGAGAAG 2368
DB 63 ----- 63
QY 2369 GGAAGCAGGTGTGGGGGACGTTCTCTGGGGAAGGACATTTATATCATGSCATGAAGGACTG 2428
DB 63 ----- 63
QY 2429 ATTTTCCAAAGCCCAAGAGAGTAGGAGGGCAAGGGCTTGGAGGTGGAGCTGGCAG 2488
DB 63 ----- 63
QY 2489 TGGGCATGCAAGCCCATTTGGGCAACATATGTTATGGAGTACAAGTCCCTTCTGTGTGACA 2548
DB 63 ----- 63
QY 2549 CCAGAAGGAAAGCGCTTGGGAATGGAAGATGATGTAGTCTCTGATGCCCGTTTAAATCAG 2608
DB 63 ----- 63
QY 2609 AAATCGAGATGAAGGGGGTGCAGTGACCCGGTTCAAACTTTTGACATGTGGTCTCTCG 2668
DB 63 ----- 63
QY 2669 GGCCTCACTGCTCAGCGGCATGGACCATCATCTGGGAATGGGATGCTAACTGGGGCTCT 2728
DB 63 ----- 63
QY 2729 CGGCAATTTGTGTGACTCTTGCAAGGTATACCTGGGTGACGCATCCAAACTGAGTTCCT 2788
DB 63 ----- 63
QY 2789 CCATCAGAGAGGTGTGACCCCAACCCCTGCCCCCGATCAGGAGGCTGGGTCTCTCTCTCT 2848

Qy	5009	TGAGCCCACTGGGGCCCAAGGACAGGACTGAGGAGAAAGGTACAGCTGGGGGGCCCT	5068
Db	331	-----	331
Qy	5069	GGGCTTAGCTGGGACACCGGGGCTTCCAGCACAGGCGTGGCCAGGCTCCTGTAAAGCTA	5128
Db	331	-----	331
Qy	5129	ACTTCCTCAACACAGGAGGAAGAGAGTGTCCCTGGGTCTGACCCATTGTGGGAGC	5188
Db	331	-----	331
Qy	5189	CATGTCGTCCAGTCCGTGTCCACAGGAGATCGACAGCTGATAGGCGAGGTGCGCGA	5248
Db	332	-----Arg--ArgValHisGlnluileAspGluValIleGlyHisValArgHis	347
Qy	5249	CCAGAGATGGGTGACACAGGCTCACATGCTTACACCACTGCGGTGATTACAGAGGTGACG	5308
Db	348	ProGluMetAlaAspGlnAlaArgMetProTyrThrAsnAlaValIleHisGluValGln	367
Qy	5309	CGTTTGGGGACATCATCCCCCTCGAGTGTGACCCATATGATCATCCGTGATCATCGAAGTA	5368
Db	368	ArgPheAlaAspIleValProThrAsnLeuProHisMetThrSerArgAspIleLysPhe	387
Qy	5369	CAGGGCTTCCGATCCCTAAGTAGGCTGGCGCCCTCTCACCCAGCTCAGCACCGC	5428
Db	388	GlnAspPhePheIleProLys-----	394
Qy	5429	ACCTGGTGATAGCCCCAGCATGGCTACTGCCAGGTGGGCCCACTCTAGGAACCTTGCCA	5488
Db	394	-----	394
Qy	5489	CCTAGTCTCAATGCCACACACTGCTGTCCCCACTTGGGTGGGGGTCCAGAGTATAG	5548
Db	394	-----	394
Qy	5549	GCAGGGCTGGCTGTCCATCCAGAGCCCCCGCTCTAGTGGGGAGACAAACAGGACCTGCC	5608
Db	394	-----	394
Qy	5609	AGATGTTGGAGACCACCGCTCCAGGAGAGGGGGCAGTGTGGTGCTCTGAGAGG	5668
Db	394	-----	394
Qy	5669	TGTGACTCGCCCTGCTGTGGGTGGAGAGGTACTGTGAGCTTCTCGGGCGCAGGAC	5728
Db	394	-----	394
Qy	5729	TAGTTGACAGAGTCCAGCTGTGTCCAGGCAGTGTGTCCCCGGTGTGTTGGTGGCAG	5788
Db	394	-----	394
Qy	5789	GGGTCCAGCATCTTAGAGTCCAGTCCCCACTCTCACCCCTGCATCTCTGCCCCAGGGAAC	5848
Db	395	-----Gly--Th	396
Qy	5849	GACACTCATACCAACCTGTTCATCGGTGCTGAAGGATGAGCGCTCTGGGAGAAGCCCTT	5908
Db	396	rThrLeuIleProAsnLeuSerValLeuLysAspGluThrValTrpGluLysProLe	416
Qy	5909	CCGCTTCCACCCGAACTTCTCGATGCCAGGGCCACTTTGTGAAGCCGAGGACCTT	5968
Db	416	uArgPheTyrProGluHisPheLeuAspAlaGlnGlyHisPheValLysHisGluAlaPh	436
Qy	5969	CCTGCCCTTCTCAGCAGGTGCTGTGGGGAGCCGGCTCCCTGTCCCTTCCGTGGAGTC	6028
Db	436	eMetProPhe--Ser-----	440
Qy	6029	TTGCAGGGGTATCACCCAGGAGCCAGGCTCACTGACGCCCTCCCTCCCCACAGGCGCG	6088
Db	441	-----AlaGlyArg	443
Qy	6089	CGTGCATGCTCGGGAGACCCCTGGCCCGCATGGAGCTTCTCTCTTCTTCACTCCCTG	6148

Db 444 ArgSerCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrCysLeu 463
 QY 6149 CTGCAGCACATTCTAGCTTCTCCGTGGCGCGGACAGCCCGCGGACCTCTCGGTGC 6208
 Db 464 LeuGlnArgPheSerValProAspGlyGlnProArgProSerAspTyrGlyLe 483
 QY 6209 GTTCAGCTTTCTGTGACGCCCATCCCTACGAGCTTTGTGCTGTGCCCGC 6259
 Db 484 TyrThrMetProValThrProGluProTyrGlnLeuCysAlaValAlaArg 500
 RESULT 12
 Q9DBJ5 PRELIMINARY; PRT; 500 AA.
 ID Q9DBJ5 AC
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE 1300006E06Rik protein.
 GN CYP2D26 OR 1300006E06RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Havaehizaki Y.
 RT "Functional annotation of a full-length mouse cDNA collection."
 RC Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DE EMBL; AK004915; BAB23666.1; -;
 DR HSSP; P00179; 1DT6.
 DR MGD; MGI:19231529; Cyp2d26.
 DR GO; GO:0004437; F:monooxygenase activity; IEA.
 DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. .; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome P450.
 DR InterPro; IPR008069; EP450_CYP2D.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01686; EP450ICYP2D.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 500 AA; 56949 MW; F4C9A03E04C8752D CRC64;
 Alignment Scores:
 Pred. No.: 3,51e-63 Length: 500
 Score: 1162.50 Matches: 348
 Percent Similarity: 28.72% Conservative: 55
 Best Local Similarity: 24.80% Mismatches: 94
 Query Match: 6.21% Indels: 906
 DB: 11 Gaps: 10
 US-09-820-788A-3 (1-10278) x Q9DBJ5 (1-500)

QY 3833 GGCAAAAGCGGAACTGGGAAGCGGGGGA CGGAGAAGCAACCCCTTACCGCATCTC 3892
Db 168 rAlaGlnAla----- 171
QY 3893 CCCACCCCGAGGAGCCCTTTCCGCCCAACGCTCTTGGACAAAGCCGTGAGCAACGT 3952
Db 172 -----GlyGlnProIleAsnProMetLeuAsnLysSerThrCysAsnVa 188
QY 3953 GATCGCTCCTCACCTGGCGGCGCTTCGAGTACGACGACCTCGCTTCTCAGGCT 4012
Db 188 lIleAlaSerLeuLlePheAlaArgArgPheGluTyrGluAspProPheLeuIleArgMe 208
QY 4013 GCTGGACCTAGCTCAGGAGGAGCTGAAGAGGAGTGGGCTTCTCGCGAGGTGGAG 4072
Db 208 tLeuLysValLeuGluGlnSerLeuThrGluValSerGlyLeuIleProGlu----- 225
QY 4073 CGAGACCGAGGAGTCTCTGACGGGCGAGCTCTCAGAGGTGCGGGGCTGAGTGGGG 4132
Db 225 ----- 225
QY 4133 CCTCCGAAGGCGAGGATTTGCATAGATGGTTTGGGAAAGGACATTCCAGGAGACCCAC 4192
Db 225 ----- 225
QY 4193 TGTAAAGAGGCGCTGGAGGAGGAGGACATCTCAGACATGTCGTGGGAGAGTGTGCC 4252
Db 225 ----- 225
QY 4253 CGGGTCAGGGGACACAGAGAGGCCAAGACTCTGTACCCCGTCCAGTTGGAGATT 4312
Db 225 ----- 225
QY 4313 CGATTTTAGGTTTCTCTGGGCAAGAGAGAGGCTGGAGGTGCACTTGGGGAGGA 4372
Db 225 ----- 225
QY 4373 CTTGGTGAAGTCAAGTGAAGGACAGGCGCCCTGGGTCTACCTGGAGATGCTGGGGC 4432
Db 225 ----- 225
QY 4433 CTGAGACTTGTCTCAGGTGAACGAGAGCAAGAGGGATTTGAGACCCCGTCTGTCTGT 4492
Db 225 ----- 225
QY 4493 GTAGGTGCTGAATGCTGCTCCCGTCTCTGCACATCCAGCGCTGGCTGGCAAGTCTCT 4552
Db 226 ---ValLeuAsnAlaPheProIleLeuLeuArgIleProArgLeuAlaAspLysAlaLe 244
QY 4553 AGCTTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGCTGCTAACTAGCACAGGATGAC 4612
Db 244 uGlnGlyGlnLysSerPheIleAlaIleLeuAspAsnLeuLeuThrGluAsnArgThrTh 264
QY 4613 CTGGGACCCAGGCCACCGACCTGAGTCTGAGGCTTCTCGGCAAGAGAGAA 4672
Db 264 rTrpAspProValGlnAlaProArgAsnLeuThrAspAlaPheLeuAlaIleGluLyl 284
QY 4673 GGTGAGTGGCTGCCACCGTGGGGGCAAGGTGGTGGTGAACGTCCAGGAGGAAT 4732
Db 284 ----- 284
QY 4733 GAGGGAGGCTGGGCAAGGTTGGACAGTGCATCACCCGGGAGCCGCATCTGGGCTG 4792
Db 284 ----- 284
QY 4793 ACAGGTGCGAAATTGGAGTCAATTGGGGCTACCCCGTTCTATCCCTGAGTATCCTCT 4852
Db 284 ----- 284
QY 4853 CGGCCCTGCTCAGGCCAAGGAGGCCCTGAGAGCAGCTTCAATGATGAGAACTGGGCAAT 4912
Db 285 -----AlaLysGlyAsnProGluSerPheAsnAspGluAsnLeuLeuMe 300

QY 4913 AGTGTGGGTAACTGTTCTTTCGCGGGATGTGTACCACTCGACCACTGCTGCTGGGG 4972
Db 300 tValValArgAspLeuPheGlyAlaGlyMetLeuThrThrSerThrLeuSerTrpAl 320
QY 4973 CCTCTGCTCATGATCTACACCTCGATGTGTGAGCGGTGAGCCAGCTGGGGCCCAAGGCA 5032
Db 320 aLeuMetLeuMetIleLeuHisProAspValGln----- 331
QY 5033 GGGACTGAGGAGGAAGGTCATGCTGGGGGCCCTTGGGCTTAGCTGGGACACCCCGGGC 5092
Db 331 ----- 331
QY 5093 TTCCAGCACAGGCGTGGCCAGGCTCCTGTAAAGCCTAACTTCTCCAACACAGGAGGAAG 5152
Db 331 ----- 331
QY 5153 AGAGTGTCCCTCGGTGTGACCCATTGTGGGGACGCATGTCTGTCCAGTCCGTGTCCAA 5212
Db 332 ---Arg-ArgValGln 335
QY 5213 CAGGAGATCGACGCTGATAGGCGAGGTGCGGACACAGAGATGGGTACACGAGGTCTCAC 5272
Db 336 GlnGluIleAspGluValIleGlyGlnValArgHisProGluMetAlaAspGlnAlaHis 355
QY 5273 ATGCCCTACACCACTGCGCTGATTACAGAGTGTGACGCGCTTTGGGACATCATCCCTCG 5332
Db 356 MetProTyrThrAsnAlaValIleHisGluValGlnArgPheGlyAspIleValProVal 375
QY 5333 AGTGTGACCCATATACATCCCGTGCATCGAAGTACAGGGCTTCGGCATCCCTAAGGTA 5392
Db 376 AsnLeuProArgIleThrSerHisAspIleGluValGlnAspPheLeuIleProLys--- 394
QY 5393 GSCGTGGCGCCCTCTCCACCCAGCTCAGCACGACCTGTGTATAGCCCCCAGCATGGC 5452
Db 394 ----- 394
QY 5453 TACTCCAGGTGGGCCCACTCTAGGAACCTGGCCACCTAGTCTCAATGCCACCACT 5512
Db 394 ----- 394
QY 5513 GACTGTCCCACTTGGGTGGGGGTCCAGAGTATAGGAGGGCTGGCTGTCTCATCCAGA 5572
Db 394 ----- 394
QY 5573 GCCCGCGTCTAGTGGGGAGACAAACAGGACCTGCCAGAACTGTGAGGACCCAGCGCT 5632
Db 394 ----- 394
QY 5633 GCAGGAGAGGGGCGAGTGTGGGTGCTCTGTAGAGGTGTGACTGCGCCCTGCTGTGGGT 5692
Db 394 ----- 394
QY 5693 CGGAGAGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACAGAGTCCAGTGTGTG 5752
Db 394 ----- 394
QY 5753 CCAGGAGTGTGTCCCCCGTGTGTGGTGGCAGGGGTCCAGCATCTTAGAGTCCAG 5812
Db 394 ----- 394
QY 5813 TCCCCACTCTCACCTCGCATCTCTGCGCCAGGGAACGACATCATCACCACTGTCTATC 5872
Db 395 -----Gly-ThrIleLeuLeuProAsnMetSerSe 404
QY 5873 GGTGTGAAGATGAGGCGTCTGGGAGAAGCCCTTCGCTTCCACCCCGAACAATCTCT 5932
Db 404 rMetLeuLysAspGluSerValTrpGluLysProLeuArgPheHisProGluHisPheLe 424
QY 5933 GGATCCCGGCGCACTTTGTGAAGCGGAGGCTTCTGCGCTTCTCAGAGGTGCCTG 5992
Db 424 uAspAlaGlnGlyHisPheValLysProGluAlaPheMetProPhe--Ser----- 440
QY 5993 TGGGAGCGCGGCTCCCTGTGCTCCCTTCCGTGGAGTCTTGCAGGGGTATCACCAGGAGCC 6052


```
Db 440 ----- 440
QY 6053 AGGCTACTGACGCCCTCCCTCCACAGAGCCCGTGCATGCTCGGGAGCCCTG 6112
Db 441 -----AlaGlyArgSerCysLeuGlyAlaLeu 451
QY 6113 GCCCGCATGAGCTCTCTCTTCTTCACTCCCTGCTGAGCACTTCAGCTTCCTG 6172
Db 452 AlaArgMetGluLeuPheThrCysLeuLeuGlnArgPheSerVal 471
QY 6173 GCCCGCATGAGCTCTCTCTTCTTCACTCCCTGCTGAGCACTTCAGCTTCCTG 6232
Db 472 ProAspGlyGlnProGlnProSerAsnSerGlyValTyrglyLeuValAlaProSer 491
QY 6233 CCTAGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6268
Db 492 ProTyrglnLeuCysAlaValValArgAspGlnGly 503

RESULT 14
Q64530 PRELIMINARY; PRT; 504 AA.
AC Q64530;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE P45016a-m82.
OS Mus spretus (Western wild mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10096;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95349581; PubMed=7623810;
RA Sueyoshi T., Kobayashi R., Nishio K., Aida K., Moore R., Wada T.,
RA Handa H., Negishi M.;
RT "A nuclear factor (NF2d9) that binds to the male-specific P450 (Cyp
RT 2d-9) gene in mouse liver.";
RL Mol. Cell. Biol. 15:4158-4166(1995).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; U20088; AAC52246.1; -.
DR PIR; I49428; I49428.
DR HSP; P00179; I076.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0016712; F:oxidoreductase activity, acting on paired d. . .; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PRO1686; EP450ICYP2D.
DR PRINTS; PRO0385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
DR Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 504 AA; 57191 MW; BD1226701778991D CRC64;

Alignment Scores:
Pred. No.: 3,06e-61 Length: 504
Score: 1131.00 Matches: 345
Percent Similarity: 28.60% Conservative: 54
Best Local Similarity: 24.73% Mismatches: 93
Query Match: 6.04% Indels: 903
Db: 11 Gaps: 12

US-09-820-788A-3 (1-10278) x Q64530 (1-504)
QY 2093 CTGGTGCCTCGCCATGATAGTGGCCATCTTCCTGCTCCTGCTGAGCTGATGACCGG 2152
Db 9 LeuTrpProValAlaIlePheThrValIlePheIleLeuValAspLeuThrHisGln 28
QY 2153 CACCAACGCTGGGCTGACGCTACCGCAGGTCCTGCTGCTGCTGCTGCTGCTGCTG 2212
Db 29 ArgGlnArgTrpThrSerArgTyrProGlyProValProValProValLeuGlyAsn 48
```

```
QY 2213 CTTCGTGATGCTGGACTTCCAGAAACACACATACCTGCTTCGACAGGTGAGGAGAGGT 2272
Db 49 -LeuLeuGlnValAspLeuAspAsnMetProTyr-Ser- 60
QY 2273 CCTGGAGGCGGAGAGGTCTGAGGATGCCCCACCACACAGCAAAACATGGGTGGTGGTT 2332
Db 60 ----- 60
QY 2333 AAACACAGGCTGGATCAGAACCCAGGCTGAGAAGGGAAGACAGGTTTGGGGGACGTTCC 2392
Db 60 ----- 60
QY 2393 TGGGAGGACATTTATACATGGCATGAGGACTGGATTTTCCAAAGCCCAAGGAAGAGT 2452
Db 61 -----LeuTyr- 62
QY 2453 AGGCAAGGCGCTGGAGGTGGAGCTGGACTTGGCAGTGGGATGCAAGCCCATTTGGGCAA 2512
Db 62 ----- 62
QY 2513 CATATGTTATGGAGTACAAAGTCCCTTCTGCTGACACCAAGAAAGGCCCTTGGGAATG 2572
Db 62 ----- 62
QY 2573 GAAGATGAGTTAGTCTCTGAGTGCCGTTTAAATCAGAAATCAGAGATGAAGGGGTGCAG 2632
Db 62 ----- 62
QY 2633 TGACCCGTTTCAAACCTTTTGCACCTGTGGGTCTCTGGGCTCTCACTGCTCACCGGCATGA 2692
Db 62 ----- 62
QY 2693 CCATCATCTGGGAATGGATGCTAACTAGTGGGCTCTCTGGCAATTTTGGTGAATCTTGGCAA 2752
Db 62 ----- 62
QY 2753 GGTATACCTGGGTGACGCATCCAAACTGAGTTCTCCATCACAGAGGTGTGACCCCA 2812
Db 62 ----- 62
QY 2813 CCCCTGCCACGATCAGAGGCTGGGTCTCTCTCTCCACCTGCTCACTCTCTGTAGCC 2872
Db 62 ----- 62
QY 2873 CCGGGGTGCTCCAAAGTTCAAATAGGACTAGGACCTGTAGTCTGGGGTGTATCTGTGGCTT 2932
Db 62 ----- 62
QY 2933 GACAAGAGGCCCTGACCTCTCTCTGAGTTGCGGCGCGCTTCGGGGACGTTGTTCAGCC 2992
Db 63 -----LysLeuGlnLysArgTyrGlyAspValPheSerL 74
QY 2993 TGCAGCTGGCTGGACGCGGTGCTGCTCAATGGGCTGGCGGCGCTGCGCGAGGCGGA 3052
Db 74 euGlnMetGlyTyrLysProMetValValIleAsnGlyLeuLysAlaMetLysGluValL 94
QY 3053 TGGTGACCCGCGGAGGACACGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3112
Db 94 euLeuThrCysGlyGluAspThrAlaAspArgProProValProIlePheGluHisLeuG 114
QY 3113 GCTTGGGCGCGGTTCCTCAAGGCAAGCGGCGGTGGGGACAGAGACCGGTTTCCGTGGG 3172
Db 114 lyValLysProGlySer-Gln- 120
QY 3173 CCGGGGTGGAGTGAAGTACCGTAGCCAGCAGCGCCGACAGGGGCTGGGGTCTCTGGAAGT 3232
Db 120 ----- 120
QY 3233 GAAACAGAGATAAAGGCCAGCGAGTGGGCTGAGGACAGTGGGCCAGGAAACCACTGCAAC 3292
Db 120 ----- 120
```

QY 3293 GGGGAGGTCCGAGTCTGTGGCTGGAGAGGGGGGGGCTACTGCGCCAGACCCGCGAGAA 3352
Db 120 ----- 120
QY 3353 GCCCGGTGGCGAGGCTGATCGTGAAGTGGCGGTGGCGGGGACCCGCGCTATGTGCG 3412
Db 120 ----- 120
QY 3413 GGCTCAGTGTGGCGGAGCGGGGGGATCTTCTTGAGTGAAGGTGTGACGGGTGGG 3472
Db 120 ----- 120
QY 3473 AGAGACGAGGTGGGCCAAACCCCGCCCCAGGAGGGGAGCAATGTGGGTGACAAAGAG 3532
Db 120 ----- 120
QY 3533 TGGGCTGTGCCAGTGGACCGGGCTAGGACTGGGAGACCTTGTGGAGCGCCAGG 3592
Db 120 ----- 120
QY 3593 GTTGGAGTGGTGGCGAGGTGGGCGCAAGGCTTCATGGCAACGCCACGTGTCGCTC 3652
Db 120 ----- 120
QY 3653 CCGCCCCAGGGTGATCTGTGCGCTATGGGCGCGGTGGCGGAGAGGCGCTTC 3712
Db 121 ----- GlyValIleLeuAlaProTyrGlyProGluTrpArgGluGlnArgPhe 137
QY 3713 TCCTGTCCACCTTGGGCTGGGCAAGAGTGGTGGAGAGTGGGTGACC 3772
Db 138 SerValSerThrLeuArgAsnPheGlyLeuGlyLysSerLeuGluAspTrpValThr 157
QY 3773 GAGAGCGCCCTGCTTGTGCGCTTCGCGACCAAGCCGCTGGTGTATGGGCGAGAAG 3832
Db 158 LysGluAlaArgHisLeuCysAspAlaPhe-----ThrAlaGlnAla 171
QY 3833 GGCACAAAGCGGAACCTGGGAAGCGGGGAGCGGAGCAACCCCTTACCGCATCTC 3892
Db 172 Gly-GlnSerIleAsn----- 176
QY 3893 CCCACCCAGGAGGCCCTTTTCGCCCAACGCGCTCTGGCAAAAGCGTGAGCAACGT 3952
Db 177 -----ProAsnThrMetLeuAsnAlaValCysAsnVa 188
QY 3953 GATCGCTCCCTCACCTCGCGGCGCGCTTCGAGTACGACGACCTCGCTTCCTCAGGCT 4012
Db 188 IleAlaSerLeuIlePheAlaArgArgLeuGluTyrGluAspProTyrLeuIleArgMe 208
QY 4013 GCTGGACCTAGCTCAGAGGAGACTGAAGAGGAGTGGGCTTTCGCGGAGTGGGAG 4072
Db 208 tLeuLysValLeuLysGluCysPheThrGluIleSerGlyPhe----- 222
QY 4073 CGAGAGCCGAGAGTCTCTGACGGGCGAGCTCTGAGAGGTCCCGGGCTGACCTGGGG 4132
Db 222 ----- 222
QY 4133 CCTCCGAGGCGAGGATTTGATAGATGGGTTTGGGAAGAGACATTCCAGAGACCCAC 4192
Db 223 -----IleProGly----- 225
QY 4193 TGTGAAGAGGCGCTGGAGGAGGAGGACATCTCAGACATGTGTGGGAGAGGTGTGCC 4252
Db 225 ----- 225
QY 4253 CGGGTCAGGGGCGACCGAGAGGCCAAGACTCTGTATCCCGCTCCAGCTTGGAGATT 4312
Db 225 ----- 225
QY 4313 CGATTTTAGGTTTCTCTCTGGGCAAGGAGAGGCTGGAGCTTGGGAGGGA 4372
Db 225 ----- 225
QY 4373 CTTGGTGAGTCAAGTAAAGACAGCAGGCGCTGGGTCTACCTGGAGATGCTGGGG 4432

Db 225 ----- 225
QY 4433 CTGAGACTTGTCCAGGTGAACGACAGACACAGAGGGATTGAGACCCGCTCTGTCTGGT 4492
Db 225 ----- 225
QY 4493 GTAGGTGCTGAATGTCTCCCGTCTCTGTGCACATCCAGCGCTGGCGAAGGTCTCT 4552
Db 226 ----ValLeuAsnAlaPheProIlePheLeuArgIleProGlyLeuAlaAspMetValPh 244
QY 4553 AGCTTCCAAAAGGCTTCTGACCCAGCTGTGATGAGTGTCTTAACTGAGCAGGATGAC 4612
Db 244 eGlnGlyGlnLysSerPheMetAlaIleLeuAspAsnLeuLeuThrGluAsnArgThrTh 264
QY 4613 CTGGGACCCAGCCAGCCAGCCAGCTGAGGCTTCTTGGCAAGAGAGAGAA 4672
Db 264 rTrpAspProAspGlnProArgAsnLeuThrAlaPheLeuAlaGluIleGluLys 284
QY 4673 GGTGAGAGTGGCTGCCACCGTGGGGGCAAGGCTGGTGGTGTGAACGCTCCAGGAGAA 4732
Db 284 s----- 284
QY 4733 GAGGGAGGCTGGGCAAAAGTTGGACAGTGATCACCAGCGGAGCGGATCTGGGCTG 4792
Db 284 ----- 284
QY 4793 ACAGGTGCAGAATTGGAGGTCAATTTGGGGCTACCCGTTCTATCCCTGAGTATCTCT 4852
Db 284 ----- 284
QY 4853 CGGCGCTGTCTCAGGCCAAGGAGGCGCTGAGAGAGCTTCAATGATGAGAACCTGGCAT 4912
Db 285 -----AlaLysGlyAsnProGluSerSerPheAsnHisGluAsnLeuArgMe 300
QY 4913 AGTGTGGGTAACTTCTTCTTGGGATGTGTGACCACTCGACCACTGGCTGGGG 4972
Db 300 tValValGlyAspLeuPheThrAlaGlyMetValThrThrSerThrLeuSerTrpAl 320
QY 4973 CTTCTGTCTCATGATCTTACCTGTGAGTGTGAGCGGTGAGCCAGCTGGGGCCCAAGGCA 5032
Db 320 aLeuLeuLeuMetIleLeuHisProAspValGln----- 331
QY 5033 GCGACTGAGGGAGGAAGGTGACAGCTGGGGGCGCTTGGGCTTAGCTGGGACACCCGGGG 5092
Db 331 ----- 331
QY 5093 TTCCAGCACAGGCGTGCCAGGCTCTGTGAAGCCTAACTTCTTCCAAACAGAGGAAAG 5152
Db 331 ----- 331
QY 5153 AGAGTGTCCCTGGGTGTGACCCATTGTGGGGACGATGTCTGTCCAGTCCGTGTCCAA 5212
Db 332 -----Arg--ArgValGln 335
QY 5213 CAGGAGATCGACAGCTGATAGGCGAGGTGCGGACACAGAGATGGGTGACAGGCTCAC 5272
Db 336 GlnGluIleAspAlaValIleGlyGlnValArgHisProGluMetAlaAspGlnAlaHis 355
QY 5273 ATGCCCTACACCACTTCCCGTGTATTCAGAGGTGACGCGTGTGGGACATCATCCCGCTG 5332
Db 356 MetProTyrThrAsnAlaValIleHisGluValGlnArgPheGlyAspIleAlaProLeu 375
QY 5333 AGTGTGACCCATATGACATCCCGTGACATCGAAGTACAGGCTTCCGCACTCCCTAGGTA 5392
Db 376 AsnLeuProArgIleThrSerArgAspIleGluValGlnAspPheLeuIleProLys--- 394
QY 5393 GGCCTGGGCGCTCTCACCCAGCTCAGCACACGACCTGGTGTAGTAGCCCGCAGCATGGC 5452
Db 394 ----- 394
QY 5453 TACTGCGAGGTGGGCGCCACTTAGGAACCTTGGCCACCTAGTCTCAATGCCACCACT 5512

Db	62	-----	62
Qy	2813	CCCCTGCCCAACGATCAGSAGGCTGGGTCTCTCTCCACCTGCTCACTCTCTGGTAGCC	2872
Db	62	-----	62
Qy	2873	CCGGGGTGCTCCAAAGTTCAAATAGACTAGGACCTGTATGTCTGGGGTGATCCTGGCTT	2932
Db	62	-----	62
Qy	2933	GACAAGAGGCCCTGACCCCTCTGCTCAGTTTCGGCGCCGCTTCGGGGACGTGTTTCAGCC	2992
Db	63	-----	74
Qy	2993	TGCAGCTGGCTTGGACCGCGTGTCTGCTCAATGGCTGGCGCGTGGCGCGTGGCGAGCGCA	3052
Db	74	euGlnMetAlaTrpLysProIleValValIleGlnGlyLeuLysAlaMetCysGluMetL	94
Qy	3053	TGTTGACCCCGCGGAGGACACGCGCCGCTGCGCCCATCTACCAAGTCTCTGG	3112
Db	94	euLeuThrCysGlyGluAspThrAlaAspArgProValProIlePheGluTrpValLeuG	114
Qy	3113	GCTTCGGGCGCGTTCCCAAGGCAGCGCGGTGGGGGACAGAGACCGCGTTTCCGTGGG	3172
Db	114	lyValLysProGlySer-Gin-----	120
Qy	3173	CCC CGGTGGACAGTAGCCCTAGTCCCAAGCAGCGCCGACAGGGCGTGGGGTCTCTGGACGT	3232
Db	120	-----	120
Qy	3233	GAACAACAGATAAAGGCCAGCGAGTAGGGCTGAGGACAGTGGGGCCAGGAAACCACTGCAC	3292
Db	120	-----	120
Qy	3293	GGGGAGTGGGAGTCTGTGGGCTGGGAGGGGGCGGGGCTACTGCCAGACCAGCCGAGAA	3352
Db	120	-----	120
Qy	3353	GCCCGTGGCGAGGCTGATCGTTCGAAGTGGCGGTGGCGGGACCGCGCTATGCTGCG	3412
Db	120	-----	120
Qy	3413	GGCTCAGTGTGGCGGACGGCGGGATCTTCTTGAGTGGAAAGGTGGTCAGGGTGGGC	3472
Db	120	-----	120
Qy	3473	AGACACAGGTGGGGCCAAACCCCGCCAGCGAGGGGAGCAATGTGGGTGAGCAAGAG	3532
Db	120	-----	120
Qy	3533	TGGGCCCTGTGCCAGCTGGACCGGGCTAGGGACTGGCGAGACCTTGTGGAGGCGCAGG	3592
Db	120	-----	120
Qy	3593	GTTGGAGTGGTGGCGGAGGTGGGGCCAAAGGCCTTCATGGCAACGCCACGTGTCCGTC	3652
Db	120	-----	120
Qy	3653	CCGCCCCAGGGGTGATCTGTGCGGTATGGGCCCGCGTGGGCGGACAGAGCGCTTC	3712
Db	121	-----	137
Qy	3713	TCCGTGTCACCTTGGCACTTGGCGTCGCAAGTTCGCTGGAGCAGTGGGTGAC	3772
Db	138	SerValSerThrLeuArgAsnPheGlyLeuGlyLysSerLeuGluAspTrpValThr	157
Qy	3773	GAGGAGCCGCTGCTTTGTGCGCTTCGCGACCAAGCCGCTGGGTGATGGCCAGAAG	3832
Db	158	LysGluAlaArgHisLeuCys-----	168
Qy	3833	GGCAAAAGCGGGAACCTGGGAAGCGGGGACGAGAGGAACCCCTTATACCGCATCTC	3892

Db	168	raLaGlNaLa-----	171
Qy	3893	CCCAACCCCAAGAGCGCCCTTTCCGCCCAACGGCGCTCTTGAGCAAAAGCCGTGAGCAACGT	3952
Db	172	 -----GlyArgSerIleAsnProAsnThrMetLeuAsnLysSerThrCysAsnVa	188
Qy	3953	GATCGCTCTCCCTCACTCGCGGGCGCCGCTTCGAGTAGCAGCACCTCGCTTCCTCAGGCT	4012
Db	188	 lIleAlaSerLeuIlePheAlaArgArgPheGluIuTrGluAspProPheLeuIleArgMe	208
Qy	4013	GCTGGACCTAGCTCAGAGAGGAGCTGAAGGAGGAGTCCGGCTTCTCGCGCAGGTGCGGAG	4072
Db	208	 tLeuLyVaIleuGluGlnSerLeuThrGluIleSerGlyLeuIleProGlu	225
Qy	4073	CGAGAGACCGAGGAGTCTCTGCAGGGCGAGCTCTCGAGAGTCCGCGGGCTCGAGCTGGGG	4132
Db	225	-----	225
Qy	4133	CCTCCGAAGGCGAGGATTTGCATAGATGGGTTTGGGAAAGGACATTCAGGAGACCCAC	4192
Db	225	-----	225
Qy	4193	TGTAAAGAAGGCCCTGGAGGAGGGGACATCTCAGACATGGTCTGGGAGAGGTGTGCC	4252
Db	225	-----	225
Qy	4253	CGGTCAGGGGGCACGAGGAGAGGCCAAGGACTCTGTACCCCGTCCACGTTGGAGATTT	4312
Db	225	-----	225
Qy	4313	CGATTTTAGTTTCTCTCTGGGCAAGGAGAGGGTGGAGGTGGGCATCTGGGAGGGA	4372
Db	225	-----	225
Qy	4373	CTTGGTCAGGTCAGTGTGTAGGACAGCAGCGGCCCTGGTCTACTCGAGATGCTGGGC	4432
Db	225	-----	225
Qy	4433	CTGAGACTTGTCCAGGTGAACGAGACAGAGGGATTGAGACCCCGTTCTGTCTGTT	4492
Db	225	-----	225
Qy	4493	GTAGGTCTGAATGCTGTCCCGTCTCTGSCACATCCCAGCGCTGGCTGGCAAGTCTCT	4552
Db	226	 -----ValLeuAsnAlaPheProIleLeuLeuArgIleProArgLeuAlaAspLysAlaLe	244
Qy	4553	ACGCTTCCAAAAGCTTCTCTGACCCAGCTGGATGAGCTGCTAACTAGCAGCACAGGATGAC	4612
Db	244	 uGlnGlyGlnLysSerPheAsnAlaIleLeuAspAsnLeuLeuThrGluAsnArgThrTh	264
Qy	4613	CTGGGACCCAGCCAGCCACCCGAGACCTGACTGAGGCCCTCTCGCAAGAAGAGAA	4672
Db	264	 rTPAspProValGlnAlaProArgAsnLeuThrAspAlaPheLeuAlaGluIleGluLy	284
Qy	4673	GGTGAGTAGGCTGCCACCGTGGGGGGCAAGGTGGTGGTTGAACGTCCCAGGAGGAAT	4732
Db	284	8-----	284
Qy	4733	GAGGGGAGGCTGGGCAAAAGGTTGGACCATGTCATCACCCGGCAGCCGCATCTGGGCTG	4792
Db	284	-----	284
Qy	4793	ACAGGTGCAGAAATTGGAGGTCAATTGGGGGCTACCCCGTTCTATCCCTCAGTATCCTCT	4852
Db	284	-----	284
Qy	4853	CGGCCCTGCTCAGGCCAAGGGGAGCCCTGAGAGCAGCTTCAATGATGAGAACCTGGGCAT	4912
Db	285	 -----AlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuLeuI	300
Qy	4913	AGTGGTGGGTAACTGTTCTTCGCGGGATGGTGACCACTCGACCAACGCTGCCTGGGG	4972
Db	300	 eValValArgAspLeuPheGlyValGlyMetValThrThrSerThrThrLeuSerTrpAl	320

Qy	4973	CCTCCTGCTCATGATCTCTACACCTGGAGTGTGACGGTGTAGCCAGCTGGGGGCCCAAGGCA	5032
Db	320	aLeuLeuLeuMetIleLeuTyPProAspValGln-----	331
Qy	5033	GGGACTGAGGGAGGAAGGTATACAGCTGGGGGCCCTTGGGCTTAGCTGGGACACCCGGGGC	5092
Db	331	-----	331
Qy	5093	TTCCAGCACAGCGCTGCCAGGCTCTCTTAAGCCTAACTTCTCTCAACACAGAGGAAGG	5152
Db	331	-----	331
Qy	5153	AGAGTGTCCCTGGGTGCTGACCACTTGTGGGGAGCGCATGCTGTCTCAGTCCGTGTCCAA	5212
Db	332	-----Cys--ArgValGln	335
Qy	5213	CAGGATTCGACACGCTGATAGGCGCAGGTGCGGCGACAGAGATGGGTGACACAGGCTCAC	5272
Db	336	GlnGluLeuLeuAspGluValIleGlyGlnValArgHisProGluMetAlaAspGlnAlaHis	355
Qy	5273	ATGCCCTACACCACTGCCGTGATTACAGAGGTGCAGCGCTTTGGGGACATCATCCGCCCTG	5332
Db	356	MetProTyPThrAsnAlaValIleHisGluValGlnArgPheGlyAspIleValProVal	375
Qy	5333	AGTGTGACCCATATGACATCCGTGTGACATCGAAGTACAGGCTTCCGCATCCCTTAGGTA	5392
Db	376	AsnLeuProArgIleThrSerHisAspIleGluValGlnAspPheLeuIleProLys---	394
Qy	5393	GGCCTGGCGCCCTCTCACCCCACTCAGCACAGCAGCACTTGGTGATAGCCCCAGCATGGC	5452
Db	394	-----	394
Qy	5453	TACTGCCAGGTGGGCCCACTTATAGGAACCTTGGCCCACTAGTCTCAATGCCACCACT	5512
Db	394	-----	394
Qy	5513	GACTGTCCCACTTGGGTGGGGGTGCAGAGTATAGCAGGGCTGGCCTGTCTCATCCAGA	5572
Db	394	-----	394
Qy	5573	GCCCCGTCTAGTGGGGAGACAAACAGGACCTGCCAGAAATGTTGGAGGACCCAGCGCCT	5632
Db	394	-----	394
Qy	5633	GCAGGAGAGGGGGCACTGTGGGTGTCCTCTGAGAGTGTGACTGCGCCCTGTGTGGGGT	5692
Db	394	-----	394
Qy	5693	CGGAGAGGGTACTGTGGAGCTTCTCGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTG	5752
Db	394	-----	394
Qy	5753	CCAGGCAGTGTGTCCCGCTGTGTTTGGTGGCAGGGGTCCAGCATCTTAGAGTCCAG	5812
Db	394	-----	394
Qy	5813	TCCCCACTCTACCCCTGCATCTCTCTGCCAGGGAAAGACACTCATCACCAACTGTCACT	5872
Db	395	-----Gly--ThrIleLeuLeuProAsnMetSerSe	404
Qy	5873	GGTCTCGAAGATGAGGCGCTCTGGGAGAAGCCCTTCCGCTTCCACCCCGAACACTTCCT	5932
Db	404	rMetLeuIleAspGluThrValTrpGluIleProLeuArgPheHisProGluHisPheLe	424
Qy	5933	GGATGCCAGGCGCACTTGTGAAGCCGAGGCGCTTCTCGCTTCTTCAGCAGGTGCCTG	5992
Db	424	uAspAlaGlnGlyHisPheValIleProGluAlaPheMetProPhe--Ser-----	440
Qy	5993	TGGGGAGCCCGGCTCCCTGTCCCTTCCGTGGAGTCTTTGACAGGGGTATACCCAGGAGCC	6052
Db	440	-----	440

Search completed: February 25, 2004, 02:57:51
Job time : 581.102 secs

Qy	6053	AGGCTACTACAGCGCCCTCCCTCC	CCACAGCGCGCGTGCATGCTCGGGAGCCCTG	6112
Db	441	-----	-----AlaGlyArgSerCysLeuGlyGluSerLeu	451
Qy	6113	GC CGCATGAGCTTCTTCCTTTTCAC	CTCCCTGCTGCAGCACCTTCAGCTTCTCCGCTG	6172
Db	452	AlaArgMetGluLeuPheLeuPheThrCys	LeuLeuGlnArgPheSerPheSerVal	471
Qy	6173	GC CGCGGACAGCCCGCGCCAGCCACT	CTCTGCTGCTCAGCTTTCTGGTGCACCCCATCC	6232
Db	472	ProAspGlyGlnProArgProSerAsnSer	GlyValTyrGlyIleLeuValAlaProSer	491
Qy	6233	CCCTACAGCTTTGTGCTGTGCCCCGCT	AGAAATGGG	6268
Db	492	ProTyrGlnLeuCysAlaValIleArgAsp	GlnGly	503

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 02:35:21 ; Search time 15768 Seconds
(without alignments)
19464.949 Million cell updates/sec

Title: US-09-820-788A-3
Perfect score: 10278
Sequence: 1 agcctacaaagtctggga.....ccagggtcagtcgaggt 10278

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estom.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_pbg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	612.4	6.0	644	13	BU682161
2	608.6	5.9	761	12	BG698634
3	598.4	5.8	651	9	AV657559
4	590.4	5.7	3969	11	BC033022

5	558.8	5.4	1180	13	BQ067695
6	494.4	4.8	1201	13	EX402424
7	494.2	4.8	1201	13	EX358085
8	478.6	4.7	1066	13	EX423847
9	478	4.7	1201	13	EX402423
10	476	4.6	1314	11	AF289565
11	474.4	4.6	905	13	BU956614
12	467.4	4.5	2067	11	AF318322
13	466	4.5	1201	9	AL532400
14	453.4	4.4	526	10	BF567779
15	453.2	4.4	955	13	EX407975
16	452.2	4.4	1201	13	EX358084
17	451.8	4.4	467	10	BE047430
18	439	4.3	1038	9	AV697507
19	437.4	4.3	707	9	AV694695
20	433.8	4.2	996	9	AV697511
21	430.4	4.2	829	28	BZ607017
22	428	4.2	1201	13	EX343659
23	425.6	4.1	712	9	AV691355
24	413.6	4.0	446	9	AV698784
25	412.2	4.0	432	9	AV694378
26	412.2	4.0	434	9	AV697704
27	409.4	4.0	428	10	AW522657
28	407.4	4.0	952	28	BZ606762
29	402.6	3.9	434	9	AV691047
30	397	3.9	426	9	AV693300
31	395	3.8	678	10	BF341109
32	394.4	3.8	425	9	AV694403
33	394.4	3.8	954	13	BQ959322
34	394	3.8	428	9	AV693268
35	393.6	3.8	696	9	AV691359
36	392.8	3.8	1066	13	EX423846
37	391.4	3.8	1030	13	EX367584
38	389.8	3.8	409	10	AW532355
39	389	3.8	707	9	AV695780
40	386.8	3.8	415	9	AV689184
41	386.8	3.8	417	9	AV693266
42	384	3.7	410	9	AV691265
43	384	3.7	410	9	AV695605
44	383.2	3.7	1060	28	BZ773519
45	381	3.7	406	9	AV684718

ALIGNMENTS

RESULT 1
BU682161
LOCUS
DEFINITION
UI-CF-EC1-acb-d-19-0-UI-s1 UI-CF-EC1 Homo sapiens cDNA clone
644 bp mRNA linear EST 07-OCT-2002
UI-CF-EC1-acb-d-19-0-UI-3', mRNA sequence.

ACCESSION
BU682161
VERSION
GI:23532791
KEYWORDS
EST.

SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 644)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS
Normal, M.F., Lennon, G. and Soares, M.B.
TITLE
Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477
PubMed
8889548

COMMENT
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 17-449 >RLTR12#LTR/Retroviral (matched complement)
369-557, >RLTR12#LTR/Retroviral (matched complement) 556-644, >PTRS
(matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

Location/Qualifiers

Source

1..644
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-acb-d-19-0-UI"
/tissue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EC1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGTCTTAC.
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_LIB=UI-CF-EC1
TAG_SEQ=AAGTCTTAC"

ORIGIN

Query Match 6.0%; Score 612.4; DB 13; Length 644;
Best Local Similarity 99.5%; Pred. No. 2.1e-98;
Matches 625; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 7254 TTGCGCTGCTACTTTCAGGAGCGCTGCTTTTATTCCTTATCTGACCCACCCACATCC 7313
DB 17 TTGCGCTGCTACTTTCAGGAGCGCTGCTTTTATTCCTTATCTGACCCACCCACATCC 76
QY 7314 TCTGATTGGCCATTTTACACAGAGTGGATTGGTCCACTTACAGAGAGCTGATTGGTGC 7373
DB 77 TCTGATTGGCCATTTTACACAGAGTGGATTGGTCCACTTACAGAGAGCTGATTGGTGC 136
QY 7374 ATTTACAATCCTGAGCTAGACACAGAGTACTGATTGGTATATTACAAACCTTGAGCTA 7433
DB 137 ATTTACAATCCTGAGCTAGACACAGAGTACTGATTGGTATATTACAAACCTTGAGCTA 196
QY 7434 GACACAGTGTGTAATGGTGATTTTACAACTCCTTAGCTACACATAAAGTTGTCACAG 7493
DB 197 GACACAGTGTGTAATGGTGATTTTACAACTCCTTAGCTACACATAAAGTTGTCACAG 256
QY 7494 TCCCACTAGATTAGCTAGATAGATAGACAGAGACCTGATTGGTGGTTTACAAACC 7553
DB 257 TCCCACTAGATTAGCTAGATAGATAGACAGAGACCTGATTGGTGGTTTACAAACC 316
QY 7554 TTGAGTTAGACACAGGGTCTGACTGGTGTGTTTACAAACCTTGAGCTAGACAGAGTG 7613
DB 317 TTGAGTTAGACACAGGGTCTGACTGGTGTGTTTACAAACCTTGAGCTAGACAGAGTG 376
QY 7614 CTGATTGGTGTATTACAACTCTTTTAGCTAGAAAATAAAGTTTCCCAAGTCCCAACACAGA 7673

DB 377 CTGATTGGTGTATTTACAATCTTTTAGCTAGAAAATAAAGTTTCCCAAGTCCCAACACAGA 436
QY 7674 TTAGCTAGAT--AGAGTCTTAATTTGGTGCATGACGAACCCGGAGCTAGACACAGAGTGC 7731
DB 437 TTAGCTAGATACAGAGTGTCTAATTTGGTGCATGACGAACCCGGAGCTAGACACAGAGTGC 496
QY 7732 TGATTGGTGCATATACAAATCTCTGGCTAGACATAAAGTTTCTCAAGTCCCACTGCAC 7791
DB 497 TGATTGGTGCATATACAAATCTCTGGCTAGACATAAAGTTTCTCAAGTCCCACTGCAC 556
QY 7792 TCAGGAGCCCGACAGCTTCGCTAGTGGATCTTATGCCAGGGCCACAGGAGAGTGC 7851
DB 557 TCAGGAGCCCGACAGCTTCGCTAGTGGATCTTATGCCAGGGCCACAGGAGAGTGC 616
QY 7852 TGCTAGTCCCGACAGCCGGGACCTGTACT 7879
DB 617 TGCTAGTCCCGACAGCCGGGACCTGTACT 644
RESULT 2
BG698634
LOCUS
DEFINITION 602658476P2 NCI_CGAP_Skn3 Homo sapiens cDNA clone linear EST 07-MAY-2001
mRNA sequence.
ACCESSION BG698634
VERSION BG698634.1 GI:13966117
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 761)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10694 row: h column: 20
High quality sequence start: 22
High quality sequence stop: 646.
Location/Qualifiers
FEATURES
source
1..761
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4801603"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn3"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."
ORIGIN
Query Match 5.9%; Score 608.6; DB 12; Length 761;
Best Local Similarity 95.8%; Pred. No. 1e-97;
Matches 700; Conservative 0; Mismatches 24; Indels 7; Gaps 7;
QY 2325 GGTGGTTAAACACAGGCTGATCAGAACCCAGCTGAGAGGGAGGAGGAGTTGGG 2384
DB 22 GATGGTTAAACACAGGCTGATCAGAACCCAGCTGAGAGGGAGGAGGAGTTGGG 81
QY 2385 GACGTTCTGGGGAAGGACATTTATACATGG-CATGAAGGAGCTGGATTTTCCAAAGGCCA 2443
DB 82 GACG-TCTGGGGAAGGACATTTATACATGTCATGAAGGAGCTGGATTTTCCAAAGGCCA 140

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3969)
Strausberg.R.
Direct Submission
Submitted (21-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaabs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 34 Row: C Column: 24
This clone has the following problem: retained intron.
Location/Qualifiers
1..3969
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4828044"
/tissue_type="Testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"

Query Match 5.7%; Score 590.4; DB 11; Length 3969;
Best Local Similarity 71.5%; Pred. No. 1.6e-94;
Matches 978; Conservative 0; Mismatches 281; Indels 109; Gaps 11;

QY 7010 TGGTGGGTCGTTGCTTCTGCTCAGGAGTGTGTTCTTCTGTTGGTGTGTTGTTCTC
DB 1811 TGGTGGGTTCTTGGTCTCACTGACTCAAGATGAAGCGGTGGACCCCTCGCGGTGAGTGT
QY 7070 TACGGCTCTTAAGGCTGACAGTACGAGTGTGTTCTTCTGTTGGTGTGTTGTTCTC
DB 1751 TACAGCTCTTAAGTGGGTGTCTGGAGTGTGTTCTTCTCCGGTGGCTGTTGTTCTC
QY 7130 ACTGGCTCAGAGTGAACCTGAGTCTTC-----CAGTGTACAACCTCATAAGGCGAG
DB 1691 GCTGGGCTCAGAGTGAAGTGCAGATCTTCAATGTTGAGTGTTCAGCTCAAAAAGCAG
QY 7185 TGTGACCCCAATGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG
DB 1631 CGTGACCCCAAGAGTG---AGCAGTAGCAAGATTATTGCAAGAGCGCAAGACAAAG
QY 7245 -----GCAACCAAGTGTGCGGTGCTTCTTCTGAGGAGCCTGCTT
DB 1574 CTTCCACAGTGTGAAGGGACCCGCGGTGTCGAATGTGGCTCGGGCAGCCCGCTT
QY 7283 TTATTCCTTATCTGACCCCAACCCACATCTCTGCTGATTTGGCCCAATTTTACAGACAGTGG
DB 1514 TTATTCCTTATCTG-GCCCAACCTACATCTCTGCTGATTTG-----TAGAGCCGAG
QY 7343 ATTGCTCCACTTACAGAGCTGATTTGCTGCAATTTTACAATCCCTGAGCTAGACACAGT
DB 1464 TGGCCCTGTTTCTGAGGCGCTGATTTGGTGGCTTTTACAATCCCTGAGCTAGATACAAAG-
QY 7403 ACTGATTGGTATATTACAAACCTTTGAGCTAGACACAGAGTGTCTGAATGGTGTATTACA

1405 ---GTTCTCCAGTCCCATCAGGCGCAGCTAGATACAGAGTGTGCTGACTGGTGCACCTACA 1349
7463 ATCCCTTAGCTAGACATAAAGTTGTCCTCCAGTCCCACTAGATTAGCTAGATAGATAGA 7522
1348 AACCTTGGAGTAAACAC----- 1332
7523 CAGAGCAGCTGATTTGGTGGCTTTTACAAACCTTGGAGTTAGACACAGGCTGCTGACTGGT 7582
1331 ---AGGGTGGTGAATTTGGTGTGTTTACAAACCTTGGAGCTAGATACAGAGTGGGATTTGGT 1275
7583 TGTTTACAAACCTTGGAGCTAGACACAGAGTGTCTGATTTGGTGTATTATTACAATCTTTAGCT 7642
1274 TATTACAACTTGGAGCTAGACACAGAGTGTCTGATTTGGTGTATTATTACAATCTTTAGCT 1215
7643 AGAATAAAGTGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCC 7700
1214 AGACATAAAGACTCTCCACGCGCTCCACAGAGCAGCTAGATACAGAGTGTGACTGGTGC 1155
7701 ATGCACGAACCGGAGCTAGACACAGAGTGTCTGATTTGGTGTATATACAACTCTCTGGCTA 7760
1154 ACTCACAACCTTGGAGCTAAACACAGAGGCTGCTGATTTGGTGTATTATTACAATCCCTGAGCTA 1095
7761 GACATAAAGTGTCTCCAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCC 7820
1094 GATATAAAGACTCTCCAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCC 1035
7821 ATCTATATCCAGGCGCAGCAGCAGAGTGTCTGATTTGGTGTATATACAACTCTCTGGCTA 7880
1034 ATCTGCACTGGGCTGTCAGGCTGAGTGTCTGCTGCGAGTCTGCTGCGGCGCTGCGCATTC 975
7881 CTCAGCCCTTGGGCGAGTGTGACGCGGAGTGTGCTGCGAGCAGTGGGAGGCGACCCATCG 7940
974 CTCAGCCCTTGGGCGAGTGTGATGGGCGGCTGCGAGCAGGCGGTGCTGCTGCGG 915
7941 GGAGGCTCGGCGCTCGCAGGAGGCGCCACCGTAG---GGAGGCTTGGGCGATGGCAGGCGT 7996
914 GGAGGCTCAGGCGCGCAGCAGAGCCATGGATTGGGTAGGAGGCTCAGGCAATGGCGGCTG 855
7997 CAAGTCTTACGCTTGGCTGCGCGGAGGCTGATGAGGCTTGGGCGACAAATTCAGTGTGG 8056
854 CAGGTCCCGAGCCTTGGCGGAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGC 798
8057 TGAGGCGCGCAGGCGCAGGCGCAGTGTGCTGCGGAGCAGGCGGCTGCGAGTGTGGCGCC 8116
797 -CAGGCGCGGCGGCGCAGGCGCAGTGTGCTGCGGAGCAGGCGGCTGCGAGTGTGGCGCC 739
8117 AGGTGCTAAGCCTTCTGCTGCGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGC 8176
738 AGGTGCTAAGTCTTCTGCTGCGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGC 679
8177 GCGGCTGAGCCTTGGCGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGC 8227
678 GCGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGC 619
8228 AGGTGCGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAGGCGCAG 8287
618 CGGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 559
8288 CTTCCACAGTGTGAGGCGGCTCCACAGTGTGAGGCGCTGGGCTGAA 8335
558 CTTTGGCGCAGGCGGCGGCTCCACAGTGTGAGGCGCTGAA 511

RESULT 5
BQ067695
LOCUS
DEFINITION BQ067695 1180 bp mRNA linear EST 02-APR-2002
ACCESSION BQ067695
VERSION BQ067695.1 GI:19896741
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12828 row: d column: 07
High quality sequence start: 3
High quality sequence stop: 536.
Location/Qualifiers

FEATURES

source
1..1180
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5768406"
/lab_host="DH10B"
/clone_lib="NIH MGC 121"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: Not1;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 5.4%; Score 558.8; DB 13; Length 1180;
Best Local Similarity 98.8%; Pred. No. 7.2e-89;
Matches 563; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1004 ATTCTTAATCAGAAACATGGAGCCAGGAGCAGTGGAGATGACGACCTCAGGAGCC 1063
DB 29 ATTCTTAATCAAAACATGAGGCCAGGAGCAGTGGAGATGACGACCTCAGGAGCC 88
QY 1064 CTGGAGGATCTGTCTCAGGCTGGGCAAGGCTTCAGGCTACCACTGGGAGCTCTGG 1123
DB 89 CTGGAGGATCTGTCTCAGGCTGGGCAAGGCTTCAGGCTACCACTGGGAGCTCTGG 148
QY 1124 GAACAGCCCTGTTGCAAAACAGGAAGTCATGGCCCGCCAGAGCCAGAAATGTGGGCTGAG 1183
DB 149 GAACAGCCCTGTTGCAAAACAGGAAGTCATGGCCCGCCAGAGCCAGAAATGTGGGCTGAG 208
QY 1184 CTGGGATCCATGTGACAGCTTTGAGGCTCACCGGAGCAGCTCTGCGACGAGAGGCTCC 1243
DB 209 CTGGGATCCATGTGACAGCTTTGAGGCTCACCGGAGCAGCTCTGCGACGAGAGGCTCC 268
QY 1244 CATCCAGGAACCTCGGCAATGCTCGGAAGTGGGTACTTGTGTCGGGTCTGTATGTG 1303
DB 269 CATCCAGGAACCTCGGCAATGCTCGGAAGTGGGTACTTGTGTCGGGTCTGTATGTG 328
QY 1304 TGTGTGACTGTGTGTGAGAGAGAAATGTGTCCTGTAGTGTGAGTGTCTGTGTA 1363
DB 329 TGTGTGACTGTGTGTGAGAGAGAAATGTGTCCTGTAGTGTGAGTGTCTGTGTA 388
QY 1364 TGTGTGAATATTCTTTTGTGTGGGTGATTTTCTGCAATGTGTAATCTGTGTCCTGCAAGT 1423
DB 389 TGTGTGAATATTCTTTTGTGTGGGTGATTTTCTGCAATGTGTAATCTGTGTCCTGCAAGT 448
QY 1424 GTGAACAAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGACCAATCAGGTGTGTG 1483
DB 449 GTGAACAAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGACCAATCAGGTGTGTG 508

QY 1484 CATAGCGTCTGTGCATGTCAAGAGTGCAGAGTGAAGTGAAGGACAGGCCCATGATGCC 1543
DB 509 CATAGCGTCTGTGCATGTCAAGAGTGCAGAGTGAAGTGAAGGACAGGCCCATGATGCC 568
QY 1544 ACTCATCATCAGAGGCTCTTAAGGCCCCAGG 1573
DB 569 ACTCATCATCAGAGGCTCTTAAGGCCCCGG 598

RESULT 6

BX402424 1201 bp mRNA linear EST 13-MAY-2003
LOCUS BX402424 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1033YB14 5-PRIME, mRNA sequence.
ACCESSION BX402424
VERSION BX402424.1 GI:30618809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS1A1092B08QP1.

FEATURES

source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1033YB14"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 4.8%; Score 494.4; DB 13; Length 1201;
Best Local Similarity 76.5%; Pred. No. 2e-77;
Matches 665; Conservative 2; Mismatches 174; Indels 28; Gaps 4;
QY 7480 AAAGTTTGTCCAGTCCCACTAGATTAGCTAGATAGAGTACAGAGAGCACTGTTGG 7539
DB 68 AAAGTTTGTCCAGTCCCACTAGATTAGCTAGATAGAGTACAGAGCACTGTTGG 118
QY 7540 TCGCTTTTACAAACCTTGTAGTTAGACACAGGGTGTGACTGGTGTGTTTACAAACCTTGAG 7599
DB 119 TGCATTTCACAAACCTTGTAGCTTAACACAGGGTGTGATTGGTGTGTTTACAAACCTTGAG 178
QY 7600 CTAGACACAGAGTGTGATTGGTGTATTTTACAATCTTTTAGCTAGAAATAAAGGTTTCCC 7659
DB 179 CTAGATACAGAGTGTGATTGGTGTATTTTACAATCTCTGAGCTAGACATAAAGGTTTCTCC 238
QY 7660 AAGTCCCAACAGATTAGCTAGATAGAGTCTAATTGGTGCATGCAGAACCCGAGCTA 7719
DB 239 AAGGCCCAACAGAGAGCTAGATAGAGTGTGGTGTGCTCACTCAAAACCTTGAGCTA 298
QY 7720 GACACAGAGTGTGATTGGTGCATATACAATCTCTGCTGCTAGACATATAAAGTCTTCCAAG 7779
DB 299 GGCACAGAGTGTGATTGGTGTGTTTACAATCCCTGAGCTAGACATATAAAGCTCTCCAG 358
QY 7780 TCCCAACCTGACTCAGGAGCCAGCCAGCTTCGCTAGTGGATCTCTATGCCAGGCCACA 7839


```
DEFINITION BX423847 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CS0DM014YB01 5-PRIME, mRNA sequence.
ACCESSION BX423847
VERSION BX423847.1 GI:30659589
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1AM004ZD01QP1.
FEATURES
Source
1..1066
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM014YB01"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
ORIGIN
Query Match 4.7%; Score 478.6; DB 13; Length 1066;
Best Local Similarity 76.3%; Pred. No. 1.3e-74;
Matches 639; Conservative 5; Mismatches 174; Indels 19; Gaps 4;
QY 7514 TAGAGTAGACAGAGACATGATGGTGGTTTACAAACCTTGAGTTAGACAGAGGTGC 7573
DB 992 TGRGTAACACACAGGGTGTGATGGTGTGTTTACAAACCTTGRGCTAGATACAGAGTGC 933
QY 7574 TGACTCGTGTGTTTACAAACCTTGAGCTAGACAGAGTGTGATGGTGTATTTACAAT 7633
DB 932 CGATTGGTGTATTTACAATCTTGAGCTAGACAGAGTGTGATGGTGTATTTACAAT 873
QY 7634 CTTTGTAGCTAGAAATAAAGGTTCCCAAGTCCCAAGTCCCAAGTGTAGCTAGAT--AGAGTGTCT 7691
DB 872 CTTGTAGCTAGACATAAAGACTCTCCACGCCCTCACCAGAGCGCTAGATACAGAGTGT 813
QY 7692 AATTGTGTATGACGAACCCGGAGCTAGACAGAGTGTGATGGTGCATATACAATC 7751
DB 812 SACTGTGSACTCACAAACCTTGAGCTAAACACAGGGTGTGATGGTGTATTTACAATC 753
QY 7752 CTCTGGCTAGACATAAAGTCTTCCAGTCCCACTGACTCAGGAGCCAGCCAGCTTC 7811
DB 752 CTTGAGCTAGATATAAAGACTCTCCAGTCCCACTGAGTCCAGAGTCCAGTGGCTTC 693
QY 7812 GCCTAGTGTGATCCTATGTCAGGAGCCACAGGAGAGTGTGCTGTAGTCCCAACCCGGGCA 7871
DB 692 ACCTAGTGTGATCCTGACGTGGGGTCTCAGGTGGAGTGTGCTGCCAGTCTGTGCCGGGG 633
QY 7872 CTGTACTCTCAGCCCTTGGGAGTGGAGCGGACAGGTGCCGTGGAGCAGTGGAGGC 7931
DB 632 CTCGCAATTCCTCAGCCCTTGGGAGTGGAGTGGAGTGGGCGCGGTGGAGCAGGGGTGT 573
QY 7932 ACCCATCCGGAGGCTCGGGCCCTCGCAGGAGGCCACCGTAG----GGAGGCTTGGCAT 7987
DB 572 GCTCGTCGGGGAGGCTCAGGCGCGCACAGAGAGGCCATGGATTGGGTAGGAGGCTCAGGCAT 513
```

```
QY 7988 GGAGAGCTCAAGTCTCTGAGCCCTGCCCCCGGGAGGTGACTGAGGCTTGGCGACAATT 8047
DB 512 GCGGGCTGAGTCTCCGAGCCCTGCCCCGTGGGAAGGAGCCAGCCAGCCAGCAGAAAT 453
QY 8048 CAAAGTGTGTGAGCGCGGCGGAGGAGGAGTACTTGGGGGAGCCCGGTGCCCCCTCTGCAGC 8107
DB 452 CGAGCG-----CAGCGCGCGGTGGGCGGAGCTGCTGGGGGAGCTCAGTACACCCCTCCGAGC 397
QY 8108 TGTGCGCCAGGTGCTTAAGCCCTCTACTGCTGGGCGCAGAGGACACAGCCGCGCGCTCC 8167
DB 396 CACTGCGCCAGGTGCTTAAGTCCCCCATGTCGCCGGGCGCAGAGGCTGCGTGGCTCTCT 337
QY 8168 GAGTGCAGGGCGCGTGAAGCCCTGCCCCACCCAGAACTGGTGTGTCGCCGCGAG----- 8221
DB 336 GAGTGGGGGCGCCACCAAGCCCAAGCCCGGAACTCCAGATGGCGCGCAAGTGCAGC 277
QY 8222 ---CRAAGCCAGGTTCGCCGACAGCGCTCTCCCTCCATACCTCCCGCAAGACAGCAGC 8278
DB 276 ACACAGCGCCGCTTCGCGCTCGTGTCTCTCCCTCCACACTCCCTGCNAAGCTGAGGAGT 217
QY 8279 CGGCTCCAGCCTCCACCACTGCAGAGAGGGGCTCCCAAGTGCAGCGCTGGGCTGAA 8335
DB 216 GGGCTCCGCGCTTGGCGGAGCCAGAAAGGGGCTCCCAAGTGCAGTGGGGGAGTGA 160
```

```
RESULT 9
BX402423/c
LOCUS BX402423 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI033YB14 3-PRIME, mRNA sequence.
ACCESSION BX402423
VERSION BX402423.1 GI:30618807
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1AI009ZB08NP1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI033YB14"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 4.7%; Score 478; DB 13; Length 1201;
Best Local Similarity 76.3%; Pred. No. 1.7e-74;
Matches 632; Conservative 9; Mismatches 169; Indels 18; Gaps 4;
QY 7522 ACAGAGAGCACTGATTTGGTGGCTTTACAAACCTTGAGTTAGACAGAGGTGCTGACTGT 7581
DB 989 AAAWACAAKTTAAATTGTGCAATTTAACAAACCTTGAGCTAAACAGAGGTGCTAATTGT 930
QY 7582 GTGTTTACAAACCTTGAGCTAGACAGAGTGTGATGGTGTATTTACAATCTTTTAC 7641
```

Db 929 GTGTTAAACAACTTGGAGCTARATACAGAGTGCCATTGGTGTATTWACAATCCCTGAGC 870
Qy 7642 TAGAATAAAGGTTCCCAAGTCCCAACAGATTAGCTAGATAGAGTCTAATTCGTGCA 7701
Db 869 TAAACATAAAGGTTCTCAAGGCCCAACAGAGAGCTAGATAGAGTGTGGATTGGTGAA 810
Qy 7702 TCACGAACCCGGAGCTAGACACAGAGTGCTGATTGGTGCATATACAATCTCTCTGCTAG 7761
Db 809 CTCACAAACCTGAGCTAGGACAGAGTGCTATTGGTGTGTTTAAATCCCTGAGCTAR 750
Qy 7762 ACATAAAAGTTCTCAAGTCCCACTGACTCAGAGGCCACCGAGCTTCGCTAGTGA 7821
Db 749 ACATAAAGACTCTCAAGTCCCAACARACTCAGAGAGCTCTGGCTTCACTTAGTGA 690
Qy 7822 TCTATGCCAGGCCACAGGAGAGTGCTGCTAGTCCACACCGGACCTGTACTCC 7881
Db 689 TCCGACACCGGGGCTGCAGGTGGAGTGCCTACCAAGTCCGACAGGTGCTGTGATCC 630
Qy 7882 TCAGCCCTTGGGAGTGGACGGGACAGAGTGCCTGGAGAGTGGAGGACCCATCCGG 7941
Db 629 TCAGCCCTTGGTGTGCTGATGGAGTGGGCGCTGGAGAGGGGTGGTCTGCTGAG 570
Qy 7942 GAGGCTCGGGCTCGCAGGAGCC----CACGTTAGGAGGCTTGGGCATGSCAGGCTGC 7997
Db 569 GAGGCTCGGGCGCACAGGAGCCCTTGGAGTGGGTGGAGGCTTAGGCATGSCGGGCTGC 510
Qy 7998 AAGTCTGAGCCCTCGCGGGGAGTGACTGAGGCTGCGGACAAATCAAGTGTGT 8057
Db 509 AGGTCCTCCAGCCCTCGCGGGGAGGAGCTAAGGCTCGGTGAGAAATCGAGCG---- 454
Qy 8058 GAGCGCCGAGCGCAGCAGTACTGGGGAGCCGCTGCCCCCTCTGACGCTGCTGGCCCA 8117
Db 453 CAGCGCTGGTGGGCTGGGACTGCTGGGGAGCCACACACCTCCGAGCTGCTGGCCCG 394
Qy 8118 GGTGTAAGCCCTCACTGCTCTGG--GGCCAGAGGACCCAGCGCGGCTCCGAGTGCAGG 8176
Db 393 GGTGTAAGTCCCTCATTTGCCAGCGCGCAGCAGGCTGGCGGCTGCTCGAGTGCAGG 334
Qy 8177 GCCGCTGAGCCCTGCCACCCAGACTGGTGTGCTGGCCGCGAG-----CAAGCC 8227
Db 333 GCCGCCAAGCCAGCCGACCCGCGCACTCCAGCTGGCGGCAAGCGCGGCACACAGGCC 274
Qy 8228 AGTTTCGCGCACACGCTCTCCCTCCATACCTCCCGCAAGCAGAGCGCGGCTCCAG 8287
Db 273 CGGTTCCGCTCGCGCTCTCCCTCCACCTCCCGCAAGCTGAAGAGAGGGCTCCGG 214
Qy 8288 CTCACACAGTCCAGAGAGGGCTCCACAGTGCAGCGCTGGGCTGAA 8335
Db 213 CCTTGGCCAGCCAGAAAGGGGCTCCACAGTGCAGCGTGGGCTGAA 166

RESULT 10
AF289565
LOCUS AF289565 1914 bp mRNA linear HTC 01-JAN-2002
DEFINITION Homo sapiens clone pp6414 unknown mRNA.
ACCESSION AF289565
VERSION AF289565.1 GI:18027333
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1914)
Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T.,
Wan, D.F. and Gu, J.R.
Novel human cDNA clones with function of inhibiting cancer cell
growth
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1914)
Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T.,
Wan, D.F. and Gu, J.R.
Direct Submission
JOURNAL Submitted (17-JUL-2000) National Laboratory For Oncogenes & Related

Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai
200032, P. R. China
FEATURES
source location/Qualifiers
1. .1914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pp6414"
1043..1585
/codon_start=1
/product="unknown"
/protein_id="AAL55749.1"
/db_xref="GI:18027334"
/translation="MGLGAVEQERALLAGEAPAAOVPTGIVGSGMAGCRSLALPLGKA
ARAFGEIERSAGSALLGLDVPHPQPLARVLSPSLPGAGRACRLLRVRAKLPTQNS
SWPASFACSPGCSHLSTLSLQAEAGAGSLGPKRKLPLQCSGLGLKSSAAKVAQA
BEAPRASEGSEDCQHAVTSH"

ORIGIN

Query Match 4.6%; Score 476; DB 11; Length 1914;
Best Local Similarity 66.1%; Pred. No. 3.8e-74;
Matches 997; Conservative 0; Mismatches 250; Indels 261; Gaps 11;
Qy 6865 TGTGTCCAAAATTTGGTGGGTTCTTGGTCTCACTGACTTCAAGAAATGAAGCCGTGACCCCT 6924
Db 202 TGTGTCCGAATTTGGTGGGTTCTTGGTCTCACTGACTTCAAGAAATGAAGCCGTGACCCCT 261
Qy 6925 CACGGTAGTGTACAGTTCTTAAAGATGGTGTGTTTCAAGATTTTCTTCTGTATGTTA 6984
Db 262 CGCGGTGAGTTTACAGTTCTTAAAGCAGTGTGTCCAGAGTTTGTCTTCTACTGATGTGT 321
Qy 6985 AGAGTGTTCAGAGTTTCTTCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 7044
Db 322 GGATGTGTTCGAGATTCCTTCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 380
Qy 7045 AGCTGCAGACCTTCAACAGTGTGTTACGGCTCTTAAGCGTGCACGTACGAGTGTGTTCA 7104
Db 381 AGCTGCAGACCTTCTGTTGAGTGTACAGCTCTTAAGTGTGGTGGTGGTGGTGGTGGTGGT 440
Qy 7105 TTTCTTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 7160
Db 441 TTTCTTCCAGCGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 500
Qy 7161 -CAGTGTTCACACTCATAAAGCGAGTGTGACCCCAATGAGGAGGAGCAGCAGCAAGACT 7219
Db 501 TGAGTGTTCAGCTCATAAAGCGAGCGTGGACCCCAAGAGTGAAGTAGCAAGATTTAT 560
Qy 7220 TACTGCAACACAGCAAAAGAAATGATGGCA-----ACCAAGTTGCCGC 7260
Db 561 TCGAAGAGCAAGAACAAAGCTTCCAGATGTGGAAGAGAGACCCACAGAGATTGCCAC 620
Qy 7261 TGCTACTTCAGGAGCCCTGCTTTTATTCCTTATCTGACCCCAACCCCAACATCTCTGCTGAT 7320
Db 621 TGCTGGCTTGGGAGCCCTGCTTTTATTTCTTATTTGGCCCCCAACCCACATCTCTGGTAT 680
Qy 7321 TGGCCCATTTTACAGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 7380
Db 681 TGGT-----AGAGCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 703
Qy 7381 ATCCCTGAGCTAGACACAGAGTACTGATTGGTATATTACAAACCTTGAGCTAGACACAG 7440
Db 704 -----TGACAG 709
Qy 7441 AGTCTGAATCGTGTATTATTAATCCCTTGTAGTACATAAAGTGTGTCAGGTCCTCCAC 7500
Db 710 GGGCTGATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 769
Qy 7501 TAGATTAGCTAGATAGTACACAGAGCACTGATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 7560
Db 770 TAGATTAGCTAGAT-----ACAGAGTGTCTATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 820
Qy 7561 AGACACAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 7620

ACCESSION AF318322
VERSION AF318322.1 GI:18027735
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2067)
AUTHORS Jiang,H.Q., Zhou,X.M., Zhang,P.P., Huang,Y., Qin,W.X., Zhao,X.T.,
Wan,D.F. and Gu,J.R.
TITLE Novel human cDNA clones with function of inhibiting cancer cell
growth
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2067)
AUTHORS Jiang,H.Q., Zhou,X.M., Zhang,P.P., Huang,Y., Qin,W.X., Zhao,X.T.,
Wan,D.F. and Gu,J.R.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-2000) National Laboratory For Oncogenes & Related
Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai
200032, P. R. China
FEATURES
source
1..2067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2067
/gene="ppl1662"
1093..1701
/gene="ppl1662"
/codon_start=1
/product="unknown"
/protein_id="AAL55829.1"
/db_xref="GI:18027735"
/translation="MGLGPPALGAGRPRPRVARDLAPGAGRCRGTGSPSSAGPPA
LCLISRWALAFPRGRARDLQPAVPEPTSMGSCGASPTSTTPCSRAPSPIDHPR
AEBCERTADWAPAPAPVRDPLCEASWAPESDGVSPVLLRDCKHTNQHPVFSS
RFVSPAPVDVLLAALVGRPRFTFISSGIVNTPIGTVYLAQGL"
ORIGIN
Query Match 4.5%; Score 467.4; DB 11; Length 2067;
Best Local Similarity 72.4%; Pred. No. 1.3e-72;
Matches 694; Conservative 0; Mismatches 242; Indels 23; Gaps 6;
QY 7396 ACAGAGTACTGATTCGTATATTACAAACCTTGAGCTAGACAGAGTGTGAATGTGT 7455
DB 1891 ACAGTGTGTTGATTCGTGTTTCAATCCCTGAGCTAGATACAAAGTTTCTCCAGTCC 1832
QY 7456 ATTTACAATCCCTGAGTAGACATAAAGTTGTCCAGTCCCACTAGATTAGCTAGATA 7515
DB 1831 CCATCAGATTAGTAGACAGATTTCACACACAGTTCTCCAGGCCCCACACAGAGC 1772
QY 7516 GAGTAGACAGAGCAGCTGATTTGTCGTTTCAAAACCTTGAGTTAGACAGAGGTGCTG 7575
DB 1771 AGCTAGATACAGAGTGTGTTGTTGCTCACTACAAACCTGAGCTAGACACAGGGTGTG 1712
QY 7576 ACTGTGTGTTTACAAACCTTGAGCTAGACACAGAGTGTGTTGTTATTTACAATCT 7635
DB 1711 ATTGTGTGTTTATTAACCTTGAGCTAGATACACAGTCCGATGGTGTATTTACAATCC 1652
QY 7636 TTATCTAGATAAAGTTTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAGTCCCAAG 7693
DB 1651 CTGAGCTAGATATAAAGTTTCTCCAGGGCCCAACAGAGCAGCTAGATACAGAGTGTGA 1592
QY 7694 TTGTGTCATGACGAACCCGAGCTAGACACAGAGTGTGTTGTTGATATACAATCT 7753
DB 1591 CTGTGTCACTCAAAACCTTGAGCTAAACACAGGGTGTGTTGTTGTTTACAATCCC 1532
QY 7754 CTGGCTAGACATAAAGTTTCTCAAGTCCCACTGACTCAGGAGCCCAAGCAGCTTCCG 7813
DB 1531 TGAGCAGGACATTAAGGACTCTCCAGTCCCACTGAGCTCAGGAGCCCAAGTGGCTTCC 1472
QY 7814 CTAGTGGATCCCTATGCGAGGGCCACAGGAGAGTCCCTGCTAGTCCCAACCCGGCAC 7873

DB 1471 CTAGTGGATCCGACCGGGGCTGCAGGTGGAGCTGCCTGCAGTCCCGCGCGTGCCT 1412
QY 7874 TGTACTCTCAGCCCTTGGGAGTGGAGCGGACAGAGTGCCTGGAGAGTGGAGGGAC 7933
DB 1411 CGCACTCTCAGCCCTTGGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1352
QY 7934 CCATCCGAGGAGTTCGGGCTCGCAGGAGGAGC---CACCGTGGAGGAGTGGAGGAG 7989
DB 1351 TCGTGGGAGAGTTCGGGCTCGCAGGAGGAGC---CACCGTGGAGGAGTGGAGGAG 1292
QY 7990 CAGGCTGCAAGTCTCAGGAGTTCGGGCTCGCAGGAGGAGTTCGGGCTCGCAGGAG 8049
DB 1291 CGGCTGCAAGTCTCAGGAGTTCGGGCTCGCAGGAGGAGTTCGGGCTCGCAGGAGGAG 1232
QY 8050 AGTGTGTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8108
DB 1231 A----GCACAGCGAGGAGTTCGGGCTCGCAGGAGGAGGAGGAGGAGGAGGAGGAG 1176
QY 8109 GCTGCGCCAGGAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8168
DB 1175 GCTGCGCCAGGAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1116
QY 8169 AGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8228
DB 1115 AGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1056
QY 8229 GGTTCGCCGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8277
DB 1055 GCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 996
QY 8278 CCGGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8335
DB 995 CCGGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 937
RESULT 13
AL532400/c 1201 bp mRNA linear EST 23-MAY-2003
LOCUS AL532400 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM014YB01 5-PRIME, mRNA sequence.
ACCESSION AL532400
VERSION AL532400.2 GI:31070232
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:12795893.
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DM014CA01QPI.
FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM014YB01"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6

vector. Library was not normalized."	
ORIGIN	
Query Match	4.5%; Score 466; DB 9; Length 1201;
Best Local Similarity	76.7%; Pred. No. 2.3e-72;
Matches	638; Conservative 2; Mismatches 172; Indels 20; Gaps 5;
QY	7519 TAGACAGAGAGCTGATTTGGTCTTTACAAACCTTGAGTTAGACACAGGGTGGCTGACT 7578
Db	985 TAAACACAGGGTGTGATTTGGTCTTTACAAACCTTGAGCTAGATACAGAGTGGCGATT 926
QY	7579 GGTGTCTTTACAAACCTTGAGCTACACAGAGTCTGATTTGGTGTATTTACAACTCTTT 7638
Db	925 GGTGTATTTACAACTTGAGCTAGACAGAGTCTGATTTGGTGTATTTACAACTCTTG 866
QY	7639 AGCTAGAAATAAGGTTCCCAAGTCCCAACAGATTAGCTAGAT--AGAGTGCTAATTG 7696
Db	865 AGCTAGACATAAGACTCTCCAGGCCCTCACAGAGCAGCTAGATACAGAGTGTGACTG 806
QY	7697 GTGCATGCACGAACCCGAGCTAGACACAGAGTGTGATTTGGTGCATATCAATTCCTTG 7756
Db	805 GTGMACTCAAAACCTTGAGCTAAACACAGAGGTGCTGATTTGGTGTATTTACAACTCCCTGA 746
QY	7757 GCTAGACATAAAGTCTCCAGTCCCACTGCTAGCTAGAGAGCCAGCGACTTCGCCCTA 7816
Db	745 GCTAGATATAAAGACTCTCCAGTCCCACTGCTAGCTAGAGAGCCAGCTGGCTTCACTTA 686
QY	7817 GTGGATCTTATGCCAGGGCCACAGCAGAGTGCCTGCTAGTCCCACTCCACACCGGGCACCTGT 7876
Db	685 GTGGATCTTGCATCGGGCTCAGGTGAGTGTGCTGCCAGTCTCTGTGCCGGGGCTGCG 626
QY	7877 ACTCTCAGCCCTTGGGAGTGGACCGGACAGAGTGGCTGGAGCAGTGGGAGGAGCACCCA 7936
Db	625 ATTCTCTCAGCCCTTGGGAGTGGACCGGAGTGGGGCGCGTGGAGCAGGGGTGGTGTGCG 566
QY	7937 TCCGGAGGCTCGGGCTTCGAGGAGGCCACCGTAG----GGAGGCTTGGGAGTGGGAGTGGCAG 7992
Db	565 TCCGGAGGCTCAGGCCGACAGGAGGCCCATGGATTTGGGTAGGAGGCTCAGGCATGGCGG 506
QY	7993 GCTGCAAGTCTCAGCCCTGCCCCCGGGAGGTGACTGAGCGCTGCGCAATTCAGAT 8052
Db	505 GCTGCAAGTCTCAGCCCTGCCCCCGGGAGGAGCCAGAGGCCACCGAGAAATCGAGC 446
QY	8053 GTGGTGGAGCCCGGAGGAGCAGCAGTACTTGGGGGACCGGCTGCCCTCTGCGAGTGGCTG 8112
Db	445 G----CAGGCCGCTGGGAGCAGCAGTCTGCTGGGGAGTCAAGTACACCTCCGCGAGCAGCTG 390
QY	8113 GCCCAGTGTCTAAGCCCTCACTGCTGGGGGAGGAGGAGCAGCAGCGGCCGCTCCGAGTG 8172
Db	389 GCCCAGTGTCTAAGTCCCCCACTGCTCCGGGG--CAGCAGGGCTGGCTGGCTCTGAGTG 331
QY	8173 CAGGGCCGCTCAGCCCTGCCCCCAGAGACTGGTGTGTCGCCCGCAG-----CA 8223
Db	330 CCGGGCCCAACGAAGCCCAAGCCCACTCCAGATGTCGAGTGGAGTGGAGCAGACA 271
QY	8224 ACCCAGGTTCCCGCAGACAGCTCTCCCTCCATACCTCCCGCAGAGAGGAGCGCGCT 8283
Db	270 GCCCGGTTCCCGCTCGTGTCTCTCCCTCCACACTCCCTGCAAGCTGAGGAGTGGCT 211
QY	8284 CCAGCTTCCACAGTCCAGAGAGGGGCTCCCAAGTGGAGGCTGGGGCTGAA 8335
Db	210 CCGGCCCTGGCCAGCCAGAAAGGGGCTCCCAAGTGGAGTGGGGGAGTGA 159
RESULT 14	
BF567779/c	
LOCUS	
DEFINITION	BF567779 526 bp mRNA linear EST 12-DEC-2000
ACCESSION	UI-R-B00-ah1-c-06-0-UI.r1 UI-R-B00 Rattus norvegicus cDNA clone
VERSION	UI-R-B00-ah1-c-06-0-UI 5', mRNA sequence.
KEYWORDS	BF567779.1 GI:11677503
SOURCE	Rattus norvegicus (Norway rat)

ORGANISM	Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
REFERENCE	1 (bases 1 to 526)
AUTHORS	Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL	Genome Res. 6 (9), 791-806 (1996)
MEDLINE	97044477
PUBMED	889548
COMMENT	Contact: Soares, MB Coordinated Laboratory for Computational Genomics University of Iowa 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: bento-soares@uiowa.edu cDNA library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) This clone is also available through the I.M.A.G.E. Consortium at LLNL (info@image.llnl.gov). IMAGE ID= 1796724 Seq primer: M13 Forward.
FEATURES	Location/Qualifiers
source	1..526
	/organism="Rattus norvegicus"
	/mol_type="mRNA"
	/strain="Sprague-Dawley"
	/db_xref="taxon:10116"
	/clone="UI-R-B00-ah1-c-06-0-UI"
	/dev_stage="adult"
	/lab_host="DH10B (Life Technologies)"
	/clone_lib="UI-R-B00"
	/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The library (UI-R-B00) is a subtracted library derived from a mixture of the following tissues: thalamus, cerebellum, hypothalamus, medulla, pons, midbrain, cerebral cortex, corpus striatum and hippocampus. For a detailed description of the library from which this clone was derived, please visit our web site at rategen.eng.uiowa.edu . The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)"
ORIGIN	
Query Match	4.4%; Score 453.4; DB 10; Length 526;
Best Local Similarity	99.1%; Pred. No. 4.1e-70;
Matches	466; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY	7240 TGATGGCAACCAAGTTCGCCGTCTACTTCAGGCAGCTGCTTTTATTCCTTATCTGAC 7299
Db	477 TGATGGCAACCAAGTTCGCCGTCTACTTCAGGCAGCTGCTTTTATTCCTTATCTGAC 418
QY	7300 CCCCAACCACTCTCTGATTTGCGCCCAATTTTACAGACAGTGGATGGTCCACTTACAGA 7359
Db	417 CACCACCACTCTCTGATTTGCGCCCAATTTTACAGACAGTGGATGGTCCACTTACAGA 358
QY	7360 GAGCTGATTTGGTGCATTTACAACTCCCTGAGCTAGACACAGAGTACTGATTTGTTATTTA 7419
Db	357 GAGCTGATTTGGTGCATTTTACAATCCCTGAGCTAGACACAGAGTACTGATTTGTTATTTA 298
QY	7420 CAACCTTTGAGCTAGACACAGAGTGGTGTATTTTACAATCCCTTAGCTAGACAT 7479
Db	297 CAACCTTTGAGCTAGACACAGAGTGGTGTATTTTACAATCCCTTAGCTAGACAT 238
QY	7480 AAAGGTTGTCCCACTCCCACTAGATTAGCTAGATAGATAGACAGAGCACTGATTGG 7539
Db	237 AAAGGTTGTCCCACTCCCACTAGATTAGCTAGATAGATAGACAGAGCACTGATTGG 178
QY	7540 TGGTTTACAAACCTTGAGTTAGACACAGGGTCTGACTGGTGTGTTTACAAACCTTGAG 7599
Db	177 TGGTTTACAAACCTTGAGTTAGACACAGGGTCTGACTGGTGTGTTTACAAACCTTGAG 118

ORGANISM	Rattus norvegicus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
AUTHORS	1 (bases 1 to 526)
TITLE	Bonaldo,M.F., Lennon,G. and Soares,M.B.
JOURNAL	Normalization and subtraction: two approaches to facilitate gene discovery
MEDLINE	Genome Res. 6 (9), 791-806 (1996)
PUBMED	97044477
COMMENT	8889548 Contact: Soares, MB Coordinated Laboratory for Computational Genomics University of Iowa 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: bento-soares@uiowa.edu cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) This clone is also available through the I.M.A.G.E. Consortium at LLNL (info@image.llnl.gov). IMAGE ID= 1796724 Seq primer: M13 Forward
FEATURES	Location/Qualifiers
source	1..526 /organism="Rattus norvegicus" /mol_type="mRNA" /strain="Sprague-Dawley" /db_xref="taxon:10116" /clone="UI-R-B00-ah1-c-06-0-UI" /dev_stage="adult" /lab_host="DH10B (Life Technologies)" /clone_lib="UI-R-B00" /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The library (UI-R-B00) is a subtracted library derived from a mixture of the following tissues: thalamus, cerebellum, hypothalamus, medulla, pons, midbrain, cerebral cortex, corpus striatum and hippocampus. For a detailed description of the library from which this clone was derived, please visit our web site at ratat.eng.uiowa.edu . The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)"
ORIGIN	
Query Match	4.4%; Score 453.4; DB 10; Length 526;
Best Local Similarity	99.1%; Pred. No. 4.1e-70;
Matches	466; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY	7240 TGATGGCAACACGAGTTGGCGCTGCTACTTTCAGGAGCCCTGCTTTATTCCTTATCTGAC 7299
Db	477 TGATGGCAACGAGTTGGCGCTGCTACTTTCAGGAGCCCTGCTTTATTCCTTATCTGAC 418
QY	7300 CCCCACCCACATCTGCTGATTGGCCCAATTTTACAGACAGTGGATTGGTCCACTTACAGA 7359
Db	417 CCCCACCCACATCTGCTGATTGGCCCAATTTTACAGACAGTGGATTGGTCCACTTACAGA 358
QY	7360 GAGCTGATTGGTGCAATTTCAATCCCTGAGCTAGACACAGAGTACTGATTGGTATTATTA 7419
Db	357 GAGCTGATTGGTGCAATTTCAATCCCTGAGCTAGACACAGAGTACTGATTGGTATTATTA 298
QY	7420 CAAACCTTGAGCTAGACACAGAGTGGTGGTATTATTTACAACTCCCTTAGCTAGACAT 7479
Db	297 CAAACCTTGAGCTAGACACAGAGTGGTGGTATTATTTACAACTCCCTTAGCTAGACAT 238
QY	7480 AAAGGTTGTGCCAGTCCCACTAGATTAGTAGATAGTAGACAGAGCACTGATTGG 7539
Db	237 AAAGGTTGTGCCAGTCCCACTAGATTAGTAGATAGTAGACAGAGCACTGATTGG 178
QY	7540 TCGGTTTACAAACCTTGGAGTTAGACACAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 7599
Db	177 TCGGTTTACAAACCTTGGAGTTAGACACAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 118

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 25, 2004, 02:12:38 ; Search time 85.2513 Seconds
(without alignments)
12555.283 Million cell updates/sec

Title: US-09-820-788A-3
Perfect score: 18731
Sequence: 1 agccttaaaagtctggga.....ccagggtcagtcgaggt 10278

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2.1/USPTO spool p/US09820788/runat 24022004 141426 20026/app query.fasta_1.12174
-DB=SwissProt 42 -QPMT=fastan -SUFFIX=rs -MINMATCH=0.1 -LOOPCHI=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09820788 @CGN 1.1 143 @runat 24022004 141426 20026 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPHLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt 42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1795	9.6	497	CPD6 HUMAN	P10635 homo sapien
2	1675	8.9	497	CPDH MACFA	Q29488 macaca fasc
3	1588	8.5	497	CPDJ CALJA	O18992 callithrix
4	1296	6.9	487	CPDE BOVIN	Q01361 bos taurus
5	1280.5	6.8	500	CPD4 RAT	P13108 rattus norv
6	1273.5	6.8	500	CPDI RAT	Q64680 rattus norv
7	1266	6.8	499	CPDP PIG	O46558 sus scrofa
8	1246	6.7	499	CPDF CANFA	Q29473 canis famil
9	1198	6.4	500	CPDK MESAU	Q9QY95 mesocricetu
10	1190.5	6.4	500	CPD2 RAT	P10634 rattus norv
11	1179	6.3	500	CPDR MESAU	Q9QY96 mesocricetu
12	1177	6.3	500	CPD3 RAT	P12938 rattus norv
13	1161.5	6.2	500	CPDG CAVPO	Q64403 cavia porce
14	1146	6.1	504	CPD5 RAT	P12939 rattus norv
15	1137	6.1	504	CPD9 MOUSE	P11714 mus musculu
16	1126	6.0	504	CPDI RAT	P10633 rattus norv
17	1125	6.0	504	CPDA MOUSE	P24456 mus musculu
18	1119	6.0	500	CPDS MESAU	Q9quj1 mesocricetu

19	1034.5	5.5	505	1	CPDB MOUSE	P24457 mus musculu
20	568.5	3.0	1690	1	CA44 HUMAN	P53420 homo sapien
21	524	2.8	5703	1	MUSB HUMAN	Q9hsc84 homo sapien
22	503	2.7	2716	1	OSA DROME	Q8in94 drosophila
23	495	2.6	1690	1	CA44 HUMAN	P53420 homo sapien
24	489.5	2.6	502	1	CRJ3 RAT	P51590 rattus norv
25	485.5	2.6	2344	1	CA17 HUMAN	Q02388 homo sapien
26	481	2.6	1466	1	CA13 HUMAN	P02461 homo sapien
27	481	2.6	2944	1	CA17 HUMAN	Q02388 homo sapien
28	475	2.5	1685	1	CA54 HUMAN	P29400 homo sapien
29	469	2.5	501	1	CRJ6 MOUSE	O54750 mus musculu
30	459.5	2.5	2142	1	BAT2 HUMAN	P48634 homo sapien
31	458.5	2.5	502	1	CPJ2 HUMAN	P51589 homo sapien
32	458.5	2.5	2167	1	SHK1 RAT	Q9WV48 rattus norv
33	458	2.5	1670	1	CA34 HUMAN	Q01955 homo sapien
34	457	2.4	500	1	CPJ1 RABIT	P52786 oryctolagus
35	456.5	2.4	5262	1	MLI2 HUMAN	O14686 homo sapien
36	456	2.4	1685	1	CA54 HUMAN	P29400 homo sapien
37	455.5	2.4	1453	1	CA11 MOUSE	P11087 mus musculu
38	455.5	2.4	1763	1	CA24 ASCSU	P27393 ascaris suu
39	453.5	2.4	2161	1	SHK1 HUMAN	Q9Y566 homo sapien
40	449	2.4	1763	1	CA24 ASCSU	P27393 ascaris suu
41	448.5	2.4	1464	1	CA13 MOUSE	P08121 mus musculu
42	446.5	2.4	1669	1	CA14 MOUSE	P02463 mus musculu
43	445.5	2.4	1049	1	CA13 BOVIN	P04258 bos taurus
44	445	2.4	1806	1	CA1B HUMAN	P12107 homo sapien
45	442.5	2.4	1460	1	CA11 CANFA	Q9xsf7 canis famil

ALIGNMENTS

RESULT 1
CPD6 HUMAN
ID CPD6 HUMAN STANDARD; PRT; 497 AA.
AC P10635; Q16752;
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome P450 2D6 (EC 1.14.14.1) (CYP2D6) (P450-DB1) (Debrisoquine
DN 4-hydroxylase).
GN CYP2D6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88314109; PubMed=3410476;
RA Gonzalez F.J., Vilbois F., Hardwick J.P., McBride O.W.,
RA Nebert D.W., Gelboin H.V., Meyer U.A.;
RT "Human debrisoquine 4-hydroxylase (P450IID1): cDNA and deduced amino
RT acid sequence and assignment of the CYP2D locus to chromosome 22.";
RL Genomics 2:1174-1179(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=88122614; PubMed=3123957;
RA Gonzalez F.J., Skoda R.C., Kimura S., Umeno M., Zanger U.M.,
RA Nebert D.W., Gelboin H.V., Hardwick J.P., Meyer U.A.;
RT "Characterization of the common genetic defect in humans deficient in
RT debrisoquine metabolism.";
RL Nature 331:442-446(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90072069; PubMed=2574001;
RA Kimura S., Umeno M., Skoda R.C., Meyer U.A., Gonzalez F.J.;
RT "The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and
RT identification of the polymorphic CYP2D6 gene, a related gene, and a
RT pseudogene.";
RL Am. J. Hum. Genet. 45:889-904(1989).
RN [4]
RP VARIANT LYS-281 DEL (CYP2D6*9).
RX MEDLINE=93244880; PubMed=1844820;

RA Tyndale R., Aoyama T., Broly F., Matsunaga T., Inaba T., Kalow W.,
RA Gelboin H.V., Meyer U.A., Gonzalez F.J.;
RT "Identification of a new variant CYP2D6 allele lacking the codon
RT encoding Lys-281: possible association with the poor metabolizer
RT phenotype.";
RL Pharmacogenetics 1:26-32(1991).
RN [5]
RP VARIANTS SER-34 AND THR-486 (CYP2D6*10).
RX MEDLINE=94115362; PubMed=8287084;
RA Yokota H., Tamura S., Furuya H., Kimura S., Watanabe M., Kanazawa I.,
RA Kondo I., Gonzalez F.J.;
RT "Evidence for a new variant CYP2D6 allele CYP2D6J in a Japanese
RT population associated with lower in vivo rates of sparteine
RT metabolism.";
RL Pharmacogenetics 3:256-263(1993).
RN [6]
RP VARIANT PRO-324 (CYP2D6*7).
RX MEDLINE=95147995; PubMed=7845481;
RA Evert B., Griese E.U., Bichelbaum M.;
RT "A missense mutation in exon 6 of the CYP2D6 gene leading to a
RT histidine 324 to proline exchange is associated with the poor
RT metabolizer phenotype of sparteine.";
RL Naunyn Schmiedeberg Arch. Pharmacol. 350:434-439(1994).
RN [7]
RP VARIANT GLU-212 (CYP2D6*6B/6C).
RX MEDLINE=95172594; PubMed=7868129;
RA Daly A.K., Leathart J.B., London S.J., Idle J.R.;
RT "An inactive cytochrome P450 CYP2D6 allele containing a deletion and a
RT base substitution.";
RL Hum. Genet. 95:337-341(1995).
RN [8]
RP VARIANT IL8-107 (CYP2D6*17).
RX MEDLINE=97126511; PubMed=8971426;
RA Masimirembwa C., Persson I., Bertilsson L., Hasler J.,
RA Ingelman-Sundberg M.;
RT "A novel mutant variant of the CYP2D6 gene (CYP2D6*17) common in a
RT black African population: association with diminished debrisoquine
RT hydroxylase activity.";
RL Br. J. Clin. Pharmacol. 42:713-719(1996).
RN [9]
RP VARIANT ARG-42 (CYP2D6*12).
RX MEDLINE=96209916; PubMed=8655150;
RA Marex D., Legrand M., Sabbagh N., Lo-Guidice J.M., Boone P., Broly F.;
RT "An additional allelic variant of the CYP2D6 gene causing impaired
RT metabolism of sparteine.";
RL Hum. Genet. 97:668-670(1996).
RN [10]
RP VARIANTS.
RX MEDLINE=97385645; PubMed=9241659;
RA Marex D., Legrand M., Sabbagh N., Guidice J.M., Spire C.,
RA Lafitte J.J., Meyer U.A., Broly F.;
RT "Polymorphism of the cytochrome P450 CYP2D6 gene in a European
RT population: characterization of 48 mutations and 53 alleles, their
RT frequencies and evolution.";
RL Pharmacogenetics 7:193-202(1997).
RN [11]
RP VARIANT ARG-169 (CYP2D6*14).
RX MEDLINE=99164054; PubMed=10064570;
RA Wang S.L., Lai M.D., Huang J.D.;
RT "G169R mutation diminishes the metabolic activity of CYP2D6 in
RT Chinese.";
RL Drug Metab. Dispos. 27:385-388(1999).
CC -!- FUNCTION: Responsible for the metabolism of many drugs and
CC environmental chemicals that it oxidizes. It is involved in the
CC metabolism of drugs such as antiarrhythmics, adrenoceptor
CC antagonists, and tricyclic antidepressants.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: By pregnancy.
CC -!- POLYMORPHISM: Highly polymorphic. Oxidative drug metabolism by
CC CYP2D6 is characterized by two phenotypes, the extensive
CC metabolizer (EM) and poor metabolizer (PM). Of the Caucasian

CC populations of Europe and North America, 5%-10% are of the PM
CC phenotype and are unable to metabolize the antihypertensive drug
CC debrisoquine and numerous other drugs.
CC -!- POLYMORPHISM: Allele CYP2D6*7 was also known as CYP2D6E, allele
CC CYP2D6*9 as CYP2D6C, allele CYP2D6*10 as CYP2D6J, allele CYP2D6*17
CC as CYP2D6Z.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -!- DATABASE: NAME=Cytochrome P450 Allele Nomenclature Committee;
CC NOTE=CYP2D6 alleles;
CC WWW="http://www.imm.ki.se/CYPalleles/cyp2d6.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M20403; AA52153.1; -;
CC EMBL; X08006; CA30807.1; -;
CC EMBL; M33388; AA53500.1; -;
CC PIR; S01199; O4HUD1.
CC HSSP; P00179; 1DT6.
CC Genew; HGNC:2625; CYP2D6.
CC MIM; 124030; -;
CC InterPro; IPR001128; Cytochrome P450.
CC InterPro; IPR008069; EP450_CYP2D.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR01686; EP450CYP2D.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW METAL 443
FT METAL 443
FT VARIANT 11 11 V -> M (in allele CYP2D6*35).
FT VARIANT 26 26 R -> H (in allele CYP2D6*21).
FT VARIANT 28 28 R -> C (in allele CYP2D6*22).
FT VARIANT 34 34 P -> S (in allele CYP2D6*10 and allele
FT CYP2D6*14; poor debrisoquine metabolism).
FT VARIANT 42 42 G -> R (in allele CYP2D6*12; impaired
FT metabolism of sparteine).
FT VARIANT 85 85 A -> V (in allele CYP2D6*23).
FT VARIANT 107 107 T -> I (in allele CYP2D6*17; poor
FT debrisoquine metabolism).
FT VARIANT 169 169 G -> R (in allele CYP2D6*14; poor
FT debrisoquine metabolism).
FT VARIANT 212 212 G -> E (in allele CYP2D6*6B and allele
FT CYP2D6*6C).
FT VARIANT 237 237 A -> S (in allele CYP2D6*33).
FT VARIANT 281 281 Missing (in allele CYP2D6*9).
FT VARIANT 296 296 R -> C (in allele CYP2D6*2, allele
FT CYP2D6*12, allele CYP2D6*14 and allele
FT CYP2D6*17; dbSNP:16947).
FT VARIANT 297 297 I -> L (in allele CYP2D6*24).
FT VARIANT 311 311 S -> L (in dbSNP:1800754).
FT VARIANT 324 324 H -> P (in allele CYP2D6*7; loss of
FT activity).
FT FTid=VAR_008348.


```
QY 5438 TAGCCCCAGCATGGCTACTGCCAGGTGGGCCCACTCTAGGAACCCCTGGCCACCTAGTCTCT 5497
Db 391 -----
QY 5498 CAAATGCCACCACTGACTGTCTCCCACTTGGGTGGGGGTCCAGAGATATAGGCAGGGCTG 5557
Db 391 -----
QY 5558 GCCTGTCTCATCCAGAGCCCCCGTCTAGTGGGGAGACAAACAGGACCTGCCAGAAATGTTG 5617
Db 391 -----
QY 5618 GAGGACCCAGCGCTCGCAGGAGAGGGGCGAGTGTGGGTGCTCTGAGAGGTGTGACTGC 5677
Db 391 -----
QY 5678 GCCTGTCTGGGGTGGAGAGGGTACTGTGGAGCTTCTCGGGCGGAGGACTAGTTGACA 5737
Db 391 -----
QY 5738 GAGTCCAGCTGTGTGCCAGGACGTGTGTCTCCCGTGTGTTGGTGGCAGGGGTCCCGAG 5797
Db 391 -----
QY 5798 CATCTTAGATCCAGTCCCCACTCTCACCTTCATCTCTGCCCGGAGAACGACACTCAT 5857
Db 392 -----Gly-ThrThrLeuPhe 396
QY 5858 CACCAACCTGTATCGTGTGAAGATGAGCCGCTCTGGGAGAACCCCTTCGGCTTCCA 5917
Db 396 eThrAsnLeuSerSerValLeuIysAspGluAlaValTrpGluIysProPheArgPheHi 416
QY 5918 CCCCAGAACACTCTCTGGATGCCAGGCGCACTTTGTGAAGCGGAGGCGCTTCTGCTCTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValIysProGluAlaPheLeuProH 436
QY 5978 CTCAGCAGGTGCTGTGTGGGAGCGCGCTCCCTGTCTCCCTTCGTTGAGTCTTTCAGGGG 6037
Db 436 e-Ser----- 437
QY 6038 TATCACCAGGAGCCAGGCTCACTGAGCGCCCTCCCTCCCNACAGCGCGGTGATGTC 6097
Db 438 -----AlaGlyArgArgAlaCys 443
QY 6098 CTCGGGGAGCCCTGCGCGCATGGAGCTTCTCTCTTCTACCTCCCTGCTGTCAGCAC 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPheLeuPhePheThrCysLeuLeuGlnArg 463
QY 6158 TTCACTCTCTCTGCGCGCGGACAGCCCGCGCCAGCCACTCTCTGCTGCTGTCAGCTTT 6217
Db 464 PheSerPheSerValProAlaGlyGlnProArgProSerHisHisGlyValPheAlaPhe 483
QY 6218 CTGGTGAACCCATCCCTTACGAGCTTTGTGTGTCGCCCGC 6259
Db 484 LeuValThrProSerProTyrGluLeuCysAlaValProArg 497
```

RESULT 3

```
CPDJ-CALJA STANDARD; PRT; 497 AA.
AC O18992;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D19 (EC 1.14.14.1) (CYP1D19) (P450 2D19-1).
GN CYP2D19.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callitrich.
OX NCBI_taxid=9483;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC TISSUE=Liver;
RX MEDLINE=97223367; PubMed=9056237;
RA Igarashi T.; Sakuma T.; Isogai M.; Nagata R.; Kamataki T.;
RT "Marmoset liver cytochrome P450s: study for expression and molecular
cloning of their cDNAs."
RL Arch. Biochem. Biophys. 339:85-91(1997).
CC -!- FUNCTION: RESPONSIBLE FOR THE METABOLISM OF MANY DRUGS AND
ENVIRONMENTAL CHEMICALS THAT IT OXIDIZES.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
tissues by various foreign compounds, including drugs, pesticides,
and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; D29822; BAA22155.1; -.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EP450_CYP2D.
```

```
DR Pfam; PF00667; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
```

```
DR PRINTS; PR00385; P450.
```

```
DR PROSITE; PS00086; CYTOCHROME P450; 1.
```

```
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
```

```
KW Microsome; Endoplasmic reticulum.
```

```
FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
```

```
SQ SEQUENCE 497 AA; 55911 MW; A482ABE71E4D6CAF CRC64;
```

Alignment Scores:

```
Pred. No.: 2,24e-64 Length: 497
Score: 1588.00 Matches: 428
Percent Similarity: 32.07% Conservative: 21
Best Local Similarity: 30.57% Mismatches: 48
Query Match: 8.48% Indels: 903
DB: 1 Gaps: 8
```

US-09-820-788a-3 (1-10278) x CYPDJ-CALJA (1-497)

```
QY 2078 ATGGGGCTAGAAAGCACTGGTCCCTGGCCATGATAGTGGCCATCTTCTGCTCTGGTG 2137
Db 1 MetGlyLeuAspAlaLeuValProLeuAlaValThrValAlaIlePheValLeuVal 20
QY 2138 GACCTGATGCACCGCACCAAGCTGGGTGGCGTACCTCCCGCCAGTCCCTGCCACTG 2197
Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProProGlyProMetProLeu 40
QY 2198 CCGGGCTGGGCAACCTTGTCTGATGTGGACTTCCAGAACACACATCTGCTTCGACCA 2257
Db 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProAsnSerPhe----- 58
QY 2258 GGTGAGGAGGAGGTCTCTGGAGGGCGGAGAGGTCTCTGAGGATGCCCCACACGACAAA 2317
Db 58 ----- 58
QY 2318 CATGGGTGGTGGTTAAACCAACAGCTGGATCAGAAGCCAGCTGAGAGGGGAAGCAGG 2377
Db 58 ----- 58
QY 2378 TTTGGGGAGCGTTCTCTGGGAAGGACATTTATACATGCATGAAGACTGGATTTTCCAA 2437
Db 58 ----- 58
QY 2438 AGCCCAAGGAAGTAGGTAGGCAAGGGCTGGAGGTGGAGCTGGAGTGGGCGATGC 2497
```

Db 58 ----- 58
QY 2498 AAGCCATTGGCAACATATATTATGGAGTACAAAGTCCCTTCTGCTGACACCAAGGA 2557
Db 58 ----- 58
QY 2558 AAGCCCTTGGGAATGGAAGATGAGTTAGTCTCTGAGTGCCTTTAAATACAGAAATCGAGG 2617
Db 58 ----- 58
QY 2618 ATGAAGGGGTGAGTGACCCCGTTCAAACTTTTGCAGTGTGGGTCTCTCGGGCCTCACT 2677
Db 58 ----- 58
QY 2678 GCTACCGGCATGGACCATCATCTCGGAATGGATGCTTAACTGGGGCCTCTCGGCAATTT 2737
Db 58 ----- 58
QY 2738 TGGTGACTCTTGAAGGTCTATCTGGGTGACGCATCCAACTGAGTTCTCTCCATCACAG 2797
Db 58 ----- 58
QY 2798 AAGGTGTGACCCCACTGCTGCCACGATCAGGAGGTGGGTCTCTCTTCCACCTGC 2857
Db 58 ----- 58
QY 2858 TCACTCTGTGAGCCCGGGGTGCTCCNAGTTCAATAGACTAGGACCTGTAGTCTG 2917
Db 58 ----- 58
QY 2918 GGGTGATCCTGCTTGACAAAGAGGCCCTGACCCCTCTGTCAGTTGGGGCCGCTTCG 2977
Db 59 ----- -Aen-GlnLeuArgArgPheG 66
QY 2978 GGGAGCTGTTCAGCTGCTGAGCGGCTGACCGCGGTGTCTGCTCAATGGGCTGGCGG 3037
Db 66 lYAspValPheSerLeuGlnLeuAlaTTPThrProValValLeuAenGlyLeuGluA 86
QY 3038 CCGTGGCGAGCGATGTGACCGCGGCGAGGACACGGCCGACCGCGCTGCGCCCA 3097
Db 86 laValArgGluAlaLeuValThrArgGlyGluAspThrAlaAspArgProValProI 106
QY 3098 TCTACAGGTCTCTGGCTTCGGGCGCGTTCCAAAGCGAAGCGCGGTGGGGGACAGAGA 3157
Db 106 l eTh-GluMetLeuGlyPheGlyProHisSerGlnGly-Leu----- 119
QY 3158 CCGCGTTTCCGTGGGCCCCGGGTGGACAGTGAACCGTAGCCCAAGCAGCGCCGACAGGCG 3217
Db 119 ----- 119
QY 3218 TGGGTCTGAGCGTGAACAGAGATAAAGCCAGGAGTGGGCTGAGGACAGTGGGCCA 3277
Db 119 ----- 119
QY 3278 GGAACCACTGACCGGGGAGGTGCGAGTCTGTGGGCTGGGAGGGGGCGGCTACTGC 3337
Db 119 ----- 119
QY 3338 CCAGACCCGACAGACCCGCTGGGCGAGGCTGATGCGTTCGAAAGTGGCGGTGGCGGGGAC 3397
Db 119 ----- 119
QY 3398 CCGCGCTATGTGCGGGCTCAGTGTGGGGGGGACCGGGGGGATCTTCTTGTAGTGAAG 3457
Db 119 ----- 119
QY 3458 GTGTCAGGGTGGCAGAGACGAGGTGGGGCCAAACCCCGCCAGGAGGGGACCAATG 3517
Db 119 ----- 119
QY 3518 TGGGTGAGCAAGAGTGGGCCCCGTGTGCCAGCTGGACCGGGCTAGGGACTGCGGGAGACC 3577
Db 119 ----- 119

QY 3578 TTGTGGAGCCAGCGGTTGGAGTGGTGGCGGAGGTTGGGSCCAAGGCTTCATGGCAAC 3637
Db 119 ----- 119
QY 3638 GCCACGTGTCCGTCCCGCCCCCAGGGGTGATCTGTGCGCTATGGCCCGCGTGGCGC 3697
Db 120 ----- -PheLeuAlaArgTyrGlyProAlaTrpArg 129
QY 3698 GAGCAGAGCGCTTCTCCGTGTCCACTTGCACCTTGGCCCTGGGGCAAGAAGTCTGTG 3757
Db 130 GluGlnArgArgPheSerValSerThrLeuArgAsnLeuGlyLysLysSerLeu 149
QY 3758 GAGCAGTGGGTGACCGAGGAGCGCGCTGCTTTGTGCGCTTCCGCCACCAAGCCGCTG 3817
Db 150 GluGlnTrpValThrGluGluAlaThrTyrLeuCysAlaLa----- 163
QY 3818 GGTGATGGCCAGAGGGGCACAAAGCGGAACTGGAGAGCGGGGACGAGAGAGCAACC 3877
Db 163 ----- 163
QY 3878 CTTTACCGCATCTCCCAACCCCGAGGACGCCCTTTCGCCCAACGCGCTCTTGGACAA 3937
Db 164 ----- -Phe-AlaAspHisAlaGlyArgProPheArgProAsnGlyLeuLeuAspLy 180
QY 3938 AGCGGTGAGCAACGTGATCGCTCCCTCACCTGCGGGCGCGCTTCGAGTAGGACGACCC 3997
Db 180 sAlaValSerAsnValIleAlaSerLeuThrCysArgArgArgPheGluTyrAsnAspPr 200
QY 3998 TCGCTTCTCCTAGGCTGTGGACTAGCTCAGAGAGGACTGAAGAGGAGTCCGGCTTCT 4057
Db 200 cYsLeuLeuArgLeuLeuAspLeuThrMetGluGlyLeuLysGluGluSerGlyLeuLe 220
QY 4058 CCGCGAGGTGCGAGCGAGAGACCCAGAGAGTCTCTGCGGGCGGAGCTCTCGAGAGGTGCC 4117
Db 220 uArgGlu----- 222
QY 4118 GGGGCTGAGCTGGGGCCCTCCGAAGCGGAGGATTTGCATAGATGGGTTGGGAAGGACAT 4177
Db 222 ----- 222
QY 4178 TCCAGGAGACCCCACTGTAGAAGGGCTGGAGGAGGGGGACATCTCAGACATGGTTCG 4237
Db 222 ----- 222
QY 4238 TGGGAGAGGTGTGCCCGGGTCAGGGGGCACCCAGGAGAGGCCAAGGACTCTGTATACCCCGT 4297
Db 222 ----- 222
QY 4298 CCAGTTGGAGATTTTCGATTTTAGGTTTCTCTCTGGGCAAGGAGAGAGGGTGGAGGCTG 4357
Db 222 ----- 222
QY 4358 GCATTTGGGAGGAGACTTGGTGTAGGTCAAGTGTAGGAGAGGAGGCGCTGGGTCTACCT 4417
Db 222 ----- 222
QY 4418 GGAGATGCTGGGGCTTGAGACTTGTCCAGGTGAACCGCAGAGCACAGGAGGATTTGAGAC 4477
Db 222 ----- 222
QY 4478 CCGCTTCTGTCTGGGTAGGTGCTGAATGCTGCTCCCGCTCTCTGTCACATCCCGCGCT 4537
Db 223 ----- -ValLeuAsnAlaIleProValLeuLeuArgIleProGlyLe 236
QY 4538 GGCTGGCAAGTCTTACGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGTCTGCTAAC 4597
Db 236 uAlaGlyLysValLeuArgSerGlnLysAlaPheLeuAlaGlnLeuAspGluLeuLeuTh 256
QY 4598 TGACACAGGATGACTGGGACCCAGCCACCGGAGACCTGACTGAGGCTTCTCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276

```
QY 4658 GGCAAGAGAGGAGAGAGTGGCTGCTGCCACGGTGGGGGCAAGGGTGGTGGTTGAA 4717
Db 276 uAlaGluMetGluLys----- 281
QY 4718 CGTCCCAGGAGGAATGAGGGGAGGCTGGGCAAAAGGTTGGACCAAGTGCATCACCAGGCGA 4777
Db 281 ----- 281
QY 4778 GCCGCATCTGGGCTGACAGGTGAGAAATGGAGGTCAATTTGGGGGTACCCCGTTCTATC 4837
Db 281 ----- 281
QY 4838 CCTGAGTATCTCTGGCCCTGCTCAGGCCAAGGGAGGCCCTGAGAGAGCTTCAATGA 4897
Db 282 -----ThrLysGlyAsnProGluSerSerPheAsnAs 292
QY 4898 TGAAGACCTCGCATAGTGGTGAACCTGTTCTTCCGCGGATGCTGACCACTCGAC 4957
Db 292 pGluAsnLeuHisLeuValAlaAspLeuPheSerAlaGlyMetValThrThrSerI 312
QY 4958 CACGCTGGCTGGGGCTCTCTGCTCATGATCCTACACCTGATGTGCAGCGTGAGCCAG 5017
Db 312 eThrLeuAlaTrpGlyLeuLeuMetIleLeuHisProAspValGln----- 328
QY 5018 CTGGGGCCCAAGCAGGACTGAGGAGGAGGAGGTACAGCTGGGGCCCTGGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCGGGGCTTCAGACACAGCGTGGCCAGGCTCCTGTAAAGCTTAATCTCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGGAGGAGAGAGTGTCCCTGGTGTGCTGACCCATTGTGGGGAGCATGTCTGT 5197
Db 328 ----- 328
QY 5198 CCAGTCCGTGTCAACAGAGATCGACAGCTGATAGGCGAGTGGCGGCACAGAGATG 5257
Db 329 -Arg- -ArgValGlnGlnGluIleAspValIleGlyArgValArgProGluMet 347
QY 5258 GTGACCAAGCTCACATCCCTACACACTGCGTGTATTCAGAGGTGCGAGCTTTGGG 5317
Db 348 GlyAspGlnThrFyrMetProFyrThrThrAlaValIleHisGluValGlnArgPheAla 367
QY 5318 GACATCATCCCTGAGTGTGACCATATGACATCCCTGACATCGAAGTACAGGCTTC 5377
Db 368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CGATCCCTAAGTATAGGCTGGCGCCCTCTACCCCTGCTGACCCAGCTCAGACCACTGGTGA 5437
Db 388 LeuIleProLys----- 391
QY 5438 TAGCCCCAGCATGGCTACTGCCAGGTGGGCCCACTCTAGAAACCTTGGCCACCTAGTCT 5497
Db 391 ----- 391
QY 5498 CAATGCCACCACTGACTGTCTCCCTTGGTGGGGGTCCAGAGATATAGGAGGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTTCATCCAGAGCCCCCGTCTAGTGGGGAGACAAACAGGACCTGCCAGATGTTG 5617
Db 391 ----- 391
QY 5618 GAGGACCCAGCGCTCGAGGAGAGGGGGAGTGTGGTGGCTCTGAGAGGTGTGACTGC 5677
Db 391 ----- 391
QY 5678 GCCTGCTGTGGGTGGAGAGGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACA 5737
Db 391 ----- 391
QY 5738 GAGTCCAGCTGTGTGCCAGGCAGTGTGTCTCCCGTGTCTTGTGGTGGCAGGGGTCCAG 5797
```

```
Db 391 ----- 391
QY 5798 CATCTAGAGTCCAGTCCCACCTCTCACCCTGCATCTCTGCCCCAGGAGACGACTCAT 5857
Db 392 -----Gly-ThrThrLeuPhe 396
QY 5858 CACCAACCTGTTCATCGTGTGAAGGATGAGGCGCTCTGGAGAACCCCTTCGCGCTTCCA 5917
Db 396 eThrAsnLeuSerSerValLeuLysAspGluAlaAsnTrpGluLysProPheArgPheHi 416
QY 5918 CCCCAGAACCTTCTCGATGCCAGGCGCACCTTTGTGAAGCGGAGGCTTCCTGCTTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyArgPheValLysProGluAlaPheLeuProPh 436
QY 5978 CTCAGCAGGTGCTGTGGGGAGCCCGCTCCCTGTCTCCCTTCGCTGGAGTCTTGCAGGG 6037
Db 436 e--Ser----- 437
QY 6038 TATCACCCAGGAGCCAGGCTCACTGACGCCCTCCCTCCACAGGCGCCCGTGCATGC 6097
Db 438 -----AlaGlyArgArgAlaCys 443
QY 6098 CTCGGGAGCCCTGCGCCGATGAGACTTCTCTTCTTTCACCTCCCTGCTGCAGCAC 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPhePheThrCysLeuLeuGlnArg 463
QY 6158 TTCAGCTTCTCGTGGCGCGGACAGCCCGCCCGCCAGCCACTCTCGTCTGCTGACTTT 6217
Db 464 PheSerPheSerValProAlaGlyGlnProArgProSerProHisGlyValPheAlaPhe 483
QY 6218 CTGGTGACCCCATCCCTTACGAGCTTGTGTGTGTCGCCCGC 6259
Db 484 LeuValThrProSerProTyrGluLeuCysAlaValProArg 497
```

RESULT 4

```
CPDE BOVIN STANDARD; PRT; 487 AA.
ID CPDE BOVIN AC Q01361.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D14 (EC 1.14.14.1) (CYP2D14) (Fragment).
GN CYP2D14.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93011103; PubMed=1396678;
RA Tsuneoka Y., Matsuo Y., Higuchi R., Ichikawa Y.;
RT "Characterization of the cytochrome P-450IID subfamily in bovine
RT liver. Nucleotide sequences and microheterogeneity.";
RL Eur. J. Biochem. 208:739-746(1992).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```



```
Db 154 -----Phe-AlaAspGlnAlaGlyArgProPheSerPr 164 |||||
QY 3920 CAACGGCCTCTTGACAAAGCCGTGAGCAACGTGATCGCTCCCTCACCTGGGGCGCG 3979 |||||
Db 164 oMetGlyLeuLeuAsnLysAlaValSerAsnValIleAlaSerLeuThrPheGlyCysAr 184 |||||
QY 3980 CTTTCAGTACGACGACCCCTCGCTTCCTCAGGCTGTGACCTAGCTCAGGAGGAGCTGAA 4039 |||||
Db 184 gPheGluTyrAsnAspProArgIleIleLysLeuLeuAspLeuThrGluAspGlyLeuLy 204 |||||
QY 4040 GGAGGAGTGGGGCTTCTGGCGAGGTGGCGGAGCGGAGACCGAGGAGTCTCTGCAGGCG 4099 |||||
Db 204 sGluGluPheAsnLeuValArgLys ----- 212
QY 4100 GAGCTCCTCAGAGGTGCGGGGCTGACTGGGSCCTCCGAAGGCGAGGATTTGCATAGAT 4159 |||||
Db 212 ----- 212
QY 4160 GGGTTTGGGAAAGGACATTCACGAGAGACCCCACTGTAGAAGGCGCTGGAGGAGGAGGG 4219 |||||
Db 212 ----- 212
QY 4220 ACATCTCAGACATGGTCTGGGAGAGGTGTGCCCGGGTCAGGGGGCACCGAGAGGCCA 4279 |||||
Db 212 ----- 212
QY 4280 AGGACTCTGTACCCCGCTCCACGTTGGAGATTTCGATTTTAGTGTTCCTCTGGSCAAG 4339 |||||
Db 212 ----- 212
QY 4340 GAGAGAGGGTGGAGGCTGGCACTTGGGGAGGCACTTGTGAGTCAGTGTGAAGGACAGG 4399 |||||
Db 212 ----- 212
QY 4400 CAGGCCCTGGGTCTACTGTGAGATGGCTGGGGCCTGAGACTTGTCCAGGTGAACGCAGAG 4459 |||||
Db 212 ----- 212
QY 4460 CACAGGAGGATTGAGACCCCGTTCTGTGTGTAGTGTGTAATGCTGTCCCGTCTCT 4519 |||||
Db 213 -----ValValGluAlaValProValIle 220 |||||
QY 4520 CTTGCACATCCACGCGTGGTGGCAAGTCTTACGCTTCCAAAGGCTTTCCTGACCCA 4579 |||||
Db 220 uLeuSerIleProGlyLeuAlaAlaArgValPheProAlaGlnArgAlaPheMetAlaLe 240 |||||
QY 4580 GCTGGATGAGTCTCACTGAGCACAGGATGACTGGGACCCAGCCAGCCACCCCGAGA 4639 |||||
Db 240 uIleAspGlyLeuIleAlaGluGlnLysMetThrArgAspProThrGlnProProArgHi 260 |||||
QY 4640 CTTGACTGAGGCTTCTTGGCAAGAGGAGAGGTGAGAGTGGCTGCCACGGTGGGGGG 4699 |||||
Db 260 sLeuThrAspAlaPheLeu -----AspGluValLys ----- 270
QY 4700 CAAAGGTGGTGGGTGAAACGTCCAGAGGAATGAGGGGAGGTGGGCAAAAGTTGGAC 4759 |||||
Db 270 ----- 270
QY 4760 CAGTGCATCACCGGCGAGCCGCATCTGGGCTGACAGGTGCAGAAATTGGAGTCAATTGG 4819 |||||
Db 270 ----- 270
QY 4820 GGGCTACCCGTTCTATCCCTGAGTATCTCTCGGCCCTGTCTCAGGCCAAGGGAGCCC 4879 |||||
Db 271 -----GluAlaLysGlyAsnPr 276 |||||
QY 4880 TGAGAGCAGCTTCAATGATGAGAACCTCGCATAGTGGTGGGTAACCTGTCTCTGCCCG 4939 |||||
Db 276 oGluSerSerPheAsnAspGluAsnLeuArgLeuValValAlaAspLeuPheSerAlaGl 296 |||||
QY 4940 GATGGTGCACCACTCGACCAACGCTGGCGCTGGGGCTCTCTGCTCATGATCTTACACCTGGA 4999 |||||
```


QY 6080 ACAGCGCGCGTGCATGCTCGGGGAGCCCTGCGCGCATGAGCTCTTCTCTCTTC 6139
Db |||||
QY 428 AlAGIYAAGAGALAcysLeuGlyGluProLeuAlaAargmetGluLeuPheLeuPhe 447
Db |||||
QY 6140 ACCTCCCTGCTCAGCACTTACAGCTTCTCGTGGCGCGGACAGCGCCGCGCAC 6199
Db |||||
QY 448 ThrSerLeuGlnHisPheSerPheSerValProAlaGlyGlnProArgProSerGlu 467
Db |||||
QY 6200 TCTCGTGTGCTGAGCTTCTGTGTGACCCCATCCCTTACGAGCTTGTGTGCTG 6259
Db |||||
QY 468 HisGlyValPheAlaPheLeuValThrProAlaP-otyrglnLeuCysAlaValProArg 487
Db |||||
RESULT 5
CPD4_RAT
ID CPD4_RAT STANDARD; PRT; 500 AA.
AC P13108; O35107;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D4 (EC 1.14.14.1) (CYP2D4) (P450-DB4) (P450-CMF3)
DE (Debrisoquine 4-hydroxylase).
GN CYP2D4 OR CYP2D-4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=90189185; PubMed=2107330;
RA Matsumaga E., Umeno M., Gonzalez F. J.;
RT "The rat P450 IID subfamily: complete sequences of four closely
RT linked genes and evidence that gene conversions maintained sequence
RT homogeneity at the heme-binding region of the cytochrome P450 active
RT site."
RL J. Mol. Evol. 30:155-169(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=9096365; PubMed=9434752;
RA Wan J., Imaka S., Chow T., Hiroi T., Yabusaki Y., Funae Y.;
RT "Expression of four rat CYP2D isoforms in Saccharomyces cerevisiae
RT and their catalytic specificity."
RL Arch. Biochem. Biophys. 348:383-390(1997).
RN [3]
RP SEQUENCE OF 177-500 FROM N.A.
RX MEDLINE=89050091; PubMed=3190674;
RA Ishida N., Tawaragi Y., Inuzuka C., Sugita O., Kubota I.,
RA Nakazato H., Noguchi T., Sassa S.;
RT "Four species of cDNAs for cytochrome P450 isozymes immunorelated to
RT P450C-W/F encode for members of P450IID subfamily, increasing the
RT number of members within the subfamily."
RL Biochem. Biophys. Res. Commun. 156:681-688(1988).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; X52029; CAA36271.1; -.
DR EMBL; AB008425; BAA23125.1; -.
DR EMBL; M22331; AAA41052.1; -.
DR PIR; S16873; D31579.
DR HSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008069; EF450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 446 446 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT CONFLICT 473 473 T -> A (IN REF. 2).
FT CONFLICT 480 480 D -> N (IN REF. 2).
FT CONFLICT 483 483 I -> V (IN REF. 2).
SQ SEQUENCE 500 AA; 56697 MW; 505D29B2C0BB1F7E CRC64;
Alignment Scores:
Pred. No.: 1.37e-50 Length: 500
Score: 1280.50 Matches: 375
Percent Similarity: 29.38% Conservative: 39
Best Local Similarity: 26.61% Mismatches: 85
Query Match: 6.84% Indels: 910
Gaps: 12
US-09-820-788a-3 (1-10278) x CPD4_RAT (1-500)
QY 2051 AGGAGCCAGTTGCTAGTACGAGCAGCCATGGGCTAGAGCAGCTGTCCTCCCTGCCCATG 2110
Db |||||
QY 2 ArgMetProThrGlySerGlu-----LeuTrpProIleAlaIle 14
QY 2111 ATAGTGGCCATCTCTCTGCTGCTGATGACCAACCAACGCTGGGCTGCA 2170
Db |||||
QY 15 PheThrIlePheLeuLeuValAspLeuMethHisArgGlnArgTrpThrSer 34
QY 2171 CGTACCCGCGCAGGTCCCTCCCATGCTCCGGGCTGGGCAACCTTGTGCAATGGACTT 2230
Db |||||
QY 35 ArgTyrProGlyProValProTrpProValLeuGlyAsn-LeuLeuGlnIleAspPh 54
QY 2231 CCAGAACACACATACCTGCTTCGACCAAGTTCAGGAGGAGGCTCTGGAGGCGGCGAGAGG 2290
Db |||||
QY 54 eGlnAsnMetPro----- 58
QY 2291 TCCTGAGGATGCCCCACCACCAGCAACATGGTGGTAAACACACAGGCTGGATCA 2350
Db 58 ----- 58
QY 2351 GAAGCCAGGCTGAGAAAGGGAAGCAGGTTTGGGGACGTTCTGGGGAAGGACATTATA 2410
Db 58 ----- 58
QY 2411 CATGCATGAGGACTGGATTTCCTCAAGGCCAAGAGAGTAGGGCAAGGCGCTGGAGG 2470
Db |||||
QY 59 -----AlaGlyPheGln-Iys----- 63
QY 2471 TGGAGCTGGACTTGGCAGTGGGCATGCAAGCCCATTTGGGCAACATATGTTATGGAGTACA 2530
Db 63 ----- 63
QY 2531 AAGTCCCTTCTGCTGTGACACCAAGAGAAAGGCCCTTGGGAATGGAAGATGAGTTAGTCTGT 2590
Db 63 ----- 63
QY 2591 AGTCCGCTTTAAATACGAAATCGAGGATGAAGGGGTGAGTACCGCGGTTCAACCTT 2650
Db 63 ----- 63
QY 2651 TTGCACTGTGGTCTCGGGGCTCACTGCTCACCGGATGACCATCATCATCTGGGAATGGG 2710

Db 214 pThrLeuGluGluSerGlyPheLeu----- 223
QY 4091 CTGCAGGGGAGCTCTGAGAGGTGCCGGGCTGGACTGGGGCCCTCCGAAGGCGAGATT 4150
Db 223 ----- 223
QY 4151 TGCATAGATGGTGTGGGAAAGGACATTCCAGGAGNCCCACCTGTAAAGAGGGCCCTGGAG 4210
Db 223 ----- 223
QY 4211 GAGGAGGGGACATCTCAGACATGCTGCTGGAGAGGTGCTCCGGGTGAGGGGGCACCAG 4270
Db 223 ----- 223
QY 4271 GAGAGGCCAAGGACTCTGTATCCCCCGTCCACGTTGGAGATTTCGATTTTAGTCTCTCT 4330
Db 224 ----- 225
QY 4331 CTGGGCAAGGAGAGGGTGGAGGCTGGCACTTGGGGAGGACTTGGTGAGTCAAGTGT 4390
Db 225 ----- 225
QY 4391 AAGGACAGGCGCCCTGGGTCTACTGGAGATGCTGGGGCTGAGACTTGTCCAGGTG 4450
Db 225 ----- 225
QY 4451 AAGCAGAGACACAGGAGGATTGAGACCCCGTTCTGTCTGGTGTAGTGTCTGAATGCTGT 4510
Db 226 ----- 230
QY 4511 CCCGCTCTCTGCATCCAGCGCTGGTGGCAAGTCTTACGCTTCCAAAGGCTTT 4570
Db 230 eProMetLeuLeuHisLeuProGlyLeuLeuGlyLysValPheSerGlyLysLeuAlaPh 250
QY 4571 CTGACCCAGCTGAGTGAAGTCTAACTGAGCACAGGATGACTGGGACCCAGCCAGCC 4630
Db 250 eValAlaMetLeuAspGluLeuLeuThrGluHisLysValThrTrpAspProAlaGlnPr 270
QY 4631 ACCCGAGACCTGACTGAGGCTCTCTGGCAAGAGAGAGAGTGTGAGTGGTGGCCAC 4690
Db 270 oProArgAspLeuThrAspAlaPheLeuAlaGluValGluLys----- 284
QY 4691 GGTGGGGGCAAGGTGTGGTGTGAACGTCCAGGAGGATGAGGGGAGGTGGGCAAA 4750
Db 284 ----- 284
QY 4751 AGTTGGACAGTGCATCACCGCGGCGCATCTGGGCTGACAGGTGCAGAAATTGGAG 4810
Db 284 ----- 284
QY 4811 GTCATTTGGGGCTACCCCGTTCTATCCCTGAGTATCCTCTCGCCCTGCTCAGGCCAA 4870
Db 285 ----- 286
QY 4871 GGGGAGCCCTGAGACAGCTTCAATGATGAGAACCTGGCGCATAGTGGTGAACCTGTT 4930
Db 286 sGlyAsnProGluSerSerPheAsnAspGluAsnLeuArgValValAlaAspLeuPh 306
QY 4931 CCTTCGGGATGTGACCACTCGACACGCTGGCTGGGCTCTCTGCTCATGATCCT 4990
Db 306 eMetAlaGlyMetValThrSerThrThrLeuThrTrpAlaLeuLeuPheMetIleLe 326
QY 4991 ACACCTGATGTGACGCTGAGCCAGCTGGGGCCCAAGGAGGAGTGTCCAGCAGAGGCGAAG 5050
Db 326 uArgProAspValGln----- 331
QY 5051 GTACAGCTGGGGGCCCTTGGGCTTAGTGGGACACCCGGGCTTCCAGCAGAGGCGTGGC 5110
Db 331 ----- 331
QY 5111 CAGGCTCTGTAAAGCTTAATTCCTCCAACACAGGAGGAGAGTGTCCCTCGGTGC 5170

Db 331 ----- 331
QY 5171 TGACCATTTGGGGACCCATGTGTCCAGTCCGTGTCCAAACAGGAGATCGACACGTG 5230
Db 332 ----- 341
QY 5231 ATAGGAGGTGCGGCGACAGAGATGGTGCACAGGCTCACATGCTCCATCACCACCTGCC 5290
Db 342 lIleGlyGlnValArgArgProGluMetAlaAspGlnAlaArgMetProPheThrAsnAla 361
QY 5291 GTGATTCACAGGTGCAGCGCTTTGGGACATCATCCCTCTAGTGTGACCATATGACA 5350
Db 362 ValIleHisGluValGlnArgPheAlaAspIleLeuProLeuGlyValProHisLysThr 381
QY 5351 TCCGCTGATCGAAGTACAGGGCTTCGCAATCCCTAAGTAGGTAGGCTGGCGCTTCCTCA 5410
Db 382 SerArgAspIleGluValGlnGlyPheLeuIleProLys----- 394
QY 5411 CCCCAGCTCAGCACACGACACCTGGTGATAGCCCCCAGCATGGCTACTGCCAGGTGGGCCA 5470
Db 394 ----- 394
QY 5471 CTCTAGGAACCTGGCCACCTAGTCTCTCAATGCCACCACACTGACTGTCCCACCTTGGGT 5530
Db 394 ----- 394
QY 5531 GGGGGTCCAGATATAGGAGGGCTGGCTGTCTCATCCAGAGCCCCGTCTAGTGGGA 5590
Db 394 ----- 394
QY 5591 GACAAACAGGACCTGCCAGAAATGTTGGAGACCAGCGCTGCAGGAGAGGGGCGAGT 5650
Db 394 ----- 394
QY 5651 GTGGGTGCTCTGAGAGGTGTGACTGGCGCTGTGTGGGTTCGAGAGGGTACTGTGGA 5710
Db 394 ----- 394
QY 5711 GCTTCTGGGCGCAGGACTAGTTGACAGAGTCCAGTGTGTGCCAGGAGTGTGTGTCCC 5770
Db 394 ----- 394
QY 5771 CCGTGTGTTGTGGCAGGGGTCCAGCATCTTAGAGTCCAGTCCCCACTCTCACCCCTGC 5830
Db 394 ----- 394
QY 5831 ATCTCTGCCAGGGAACGACACTCATCACCAACTGTCTATCGGTGTGAAGGATGAGC 5890
Db 395 ----- 410
QY 5891 CGTCTGGAGAACCTTTCGCTTCCACCCCGAACACTTCTCTGATGCCAGGGCCACTT 5950
Db 410 rValTrpGluLysProLeuArgPheHisProGluHisPheLeuAspAlaGlnGlyAsnPh 430
QY 5951 TGTGAGCCGAGGCTTCTGCTTCTCTCAGAGGTGCTGTGGGGAGCCGGCTCCCT 6010
Db 430 eValIleHisGluAlaPheMetProPhe--Ser----- 440
QY 6011 GTCCCTTCCGTGAGTCTTGCAGGGGTATCACCCAGGAGCCAGGCTCACTGACGCCCT 6070
Db 440 ----- 440
QY 6071 CCCCTCCCCACAGCCCGCTGCATGCTCGGGAGCCCTCGGGCCGATCGAGCTCTTC 6130
Db 441 ----- 457
QY 6131 CTCTCTTCACTCCCTGCTGACACTTTCAGCTTCTCCGTGGCGCGGACACGCCCGG 6190
Db 458 LeuPhePheThrCysLeuLeuGlnArgPheSerValProAlaGlyGlnProArg 477
QY 6191 CCCAGCACCTCTCGTGTGCTGAGCTTCTGTGTAGCCCCCATCCCTACGAGCTTGTGCT 6250
Db 478 ProSerAsnTyrGlyValPheGlyAlaLeuThrThrProArgProTyrGlnLeuCysAla 497

```
QY 6251 GTGCCCCGC 6259
Db 498 SerProArg 500

RESULT 7
CPDP_PIG
ID CPDP_PIG STANDARD; PRT; 499 AA.
AC 04658;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D25 (EC 1.14.14.-) (CYP1D25) (Vitamin D(3) 25-
DE hydroxylase).
GN CYP2D25.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-56; 248-272 AND 407-429.
RC TISSUE=Liver;
RX MEDLINE=98086378; PubMed=9425298;
RA Postlind H., Axen E., Bergman T., Wikvall K.;
RT "Cloning, structure, and expression of a cDNA encoding vitamin D3 25-
RT hydroxylase.";
RL Biochem. Biophys. Res. Commun. 241:491-497(1997).
RN [2]
RP SEQUENCE OF 1-16.
RC TISSUE=Liver;
RX MEDLINE=93075023; PubMed=1445236;
RA Axen E., Bergman T., Wikvall K.;
RT "Purification and characterization of a vitamin D3 25-hydroxylase
RT from pig liver microsomes.";
RL Biochem. J. 287:725-731(1992).
CC -!- FUNCTION: CATALYZES THE FIRST STEP IN THE METABOLIC ACTIVATION OF
CC VITAMIN D(3) INTO 1-ALPHA,25-DIHYDROXYVITAMIN D(3), ITS ACTIVE,
CC HORMONAL FORM.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- TISSUE SPECIFICITY: Found in liver and kidney.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y16417; CAA76205.1; -
CC PIR; JC5819; JC5819.
CC HSP; P00179; 1DT5.
CC InterPro; IPR001128; Cytochrome_P450.
CC InterPro; IPR008069; EP450_CYP2D.
CC Pfam; PF00067; P450; 1.
CC PRINTS; PR01686; BP450ICYP2D.
CC PRINTS; PR00385; P450.
CC PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT INIT MET 0
FT METAL 445 445 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 499 AA; 56380 MW; 31C878B580561919 CRC64;

Alignment Scores:
Pred. No.: 61e-50 Length: 499
Score: 1266.00 Matches: 366
Percent Similarity: 29.96% Conservative: 51
Best Local Similarity: 26.29% Mismatches: 72
Query Match: 6.76% Indels: 903
DB: 1 Gaps: 10
```

```
US-09-820-788A-3 (1-10278) x CPDP_PIG (1-499)
QY 2102 CTGGCCATGATAGTGGCCATCTTCTGCTCTCTGCTGGACCTGATGCACGGCACCAACGC 2161
Db 11 LeuAlaLeuAlaMetValIlePheLeuLeuValAspLeuMetHisArgSerArg 30
QY 2162 TGGGCTGCACGCTACCGCCAGGTCCCTGCTCCCTGCTGGGCTGGGCAACCTTGTGTGCA 2221
Db 31 TrpAlaProArgTyProProGlyProMetProLeuProGlyLeuGlyAsn-LeuLeuG 50
QY 2222 TGTGGCTTCCAGAACACACCATCTGCTTCGACCAAGGTGGAGGAGGTCTCTGGAGGG 2281
Db 50 nValAsnPheGln----- 54
QY 2282 CGGCAGAGGTCTCTGAGGATGCCCCACACAGCAAAACATGGGTGGTGTAAACACAG 2341
Db 54 ----- 54
QY 2342 GCTGGATCAGAAGCCAGGCTGAGAAGGGAAGCAGGTTTGGGGGACGTTCTCTGGGAAG 2401
Db 54 ----- 54
QY 2402 ACATTTATATACATGGCATGAAGGACTGGATTTCACAAAGCCAAAGAGTAGGCAAGG 2461
Db 54 ----- 54
QY 2462 GCCTGGAGGTGGAGTGGACTTGGCAGTGGGCGATGCAAGCCCATTTGGGCAACATATGTTA 2521
Db 54 ----- 54
QY 2522 TGGAGTACAAAGTCCCTTCTGCTGACACCAGAAGGAAGGCCCTTGGGAATGGAAGATGAG 2581
Db 54 ----- 54
QY 2582 TTAGTCTGAGTGGCGTTTAAATCAGAAATCGAGGATGAAGGGGTGTCAGTGCACCCGT 2641
Db 54 ----- 54
QY 2642 TCAAACTTTTGCACGTGGGTCTCTGGGCTCTCACTGCTCACCGGATGGACCATCATCT 2701
Db 54 ----- 54
QY 2702 GGGAAATGGATGCTAACTGGGGCTCTCGGCAATTTTGTGTGACTCTTGCAGAGTCAATCC 2761
Db 54 ----- 54
QY 2762 TGGGTGACGCATCCAAACTGAGTTCTCCATCACAGAGGTGTGACCCACCCCTGCC 2821
Db 55 -----AspProArg----- 57
QY 2822 CACGATCAGAGGCTGGGTCTCTCTTCCACCTGCTCACTCTGCTAGCCCGGGGGTC 2881
Db 58 -----Leu-SerPhe----- 60
QY 2882 GTCCAAGGTTCAATATAGGACTTAGGACCTGTAGTCTGGGGTGTACTCTGGCTTGACAAGAG 2941
Db 60 ----- 60
QY 2942 CCCTGACCTCCCTCTGCGATTGGCGCGCTTCGGGACGCTTTCAGCTGCGAGCTGG 3001
Db 61 -----IleGlnLeuArgArgPheGlyAspValPheSerLeuGlnGlnI 76
QY 3002 CTGGACGCGCGGTGGTGTCTCAATGGGCTGGCGCGCTCGCGAGGCGATGGTGACCC 3061
Db 76 leTrpArgProValValValLeuAsnGlyLeuAlaValArgGluAlaLeuValSerH 96
QY 3062 GCGGCGAGGACACGCGCACCCCGCTGCTGCCCTGCGCCCATCTACAGGTCTCTGGGCTTGGGG 3121
Db 96 leSerHisGluThrSerAspArgProValPheIleLeuGluHisLeuGlyTyrglyp 116
QY 3122 CGCGTTTCCAGGCAAGCGGGGTGGGGACAGACCCGCTTTCGTGGGGCCCCGGGTG 3181
Db 116 roArgSer-Glu----- 119
```



```
QY 4307 AGATTTTCGATTTTAGGTTTCTCTCTGGGCAAGGAGAGAGGGTGAGGCTGGCACTTGGG 4366
Db 224 ----- 224
QY 4367 GAGGACTTGGTTCAGGTTCAGTGTAGGACAGCAGCGCCCTGGGTCTACCTGGAGATGGC 4426
Db 224 ----- 224
QY 4427 TGGGGCTTGAGACTTGTTCAGGTGAACGCAGAGCACAGGAGGGAATTGAGACCCCGTTCTG 4486
Db 224 ----- 224
QY 4487 TCTGTGTAGGTCTGAATGCTGTCTCCCTCTCTGTCGACATCCACGCGCTGGTGGCAA 4546
Db 225 -----AlaLeuAsnSerIleProValLeuLeuHisIleProGlyLeuAlaSerIly 241
QY 4547 GGTTCCTACGCTTCCAAAAGCCTTCTGACCCAGCTGGATGAGCTGCTAACTAGCACAG 4606
Db 241 sValPheSerAlaGlnIlyeAlaIleIleThrLeuThrAenGluMetIleGlnGluHisAr 261
QY 4607 GATGACTTGGGACCCAGCCAGCCACCCGAGACCTGACTGAGCGCTTCTCTGCGCAAGAA 4666
Db 261 glysThrArgAspProThrGlnProArgHisLeuIleAspAlaPheValAspGluIrl 281
QY 4667 GGAGAAGGTGAGAGTGGCTGCCACGGTGGGGGCAAGGGTGGTGGTGAACGTCCACGG 4726
Db 281 eGluIlyS----- 283
QY 4727 AGGAATGAGGGAGGCTGGGCAAAAGTTGGACCATGTCATCACCCGGCGAGCCGATCT 4786
Db 283 ----- 283
QY 4787 GGGCTGACAGGTGCAGAAATTGGAGGTCAATTGGGGCTACCCGTTCTATCCCTCAGTA 4846
Db 283 ----- 283
QY 4847 TCCTCTCGCCCTGTCTCAGGCAAGGGAGCCCTGAGACGACTCAATGATGAGAACT 4906
Db 284 -----AlaIlySgIyAsnProIlySThrSerPheAsnGluGluAsnIle 297
QY 4907 GGCATAGTGGTGGTAACTTCTTCCGGGATGGTGACCACTCGACCACTGGC 4966
Db 297 yCysMetValThrSerAspPheIleAlaGlyMetValSerThrSerIleThrLeuth 317
QY 4967 CTGGGCGCTCTCTGCTCATGATCTACACCTGATGTGACGCTGAGCGCTGAGCCCGAGTGGGCC 5026
Db 317 rTrpAlaLeuLeuLeuMetIleLeuHisProAspValGln----- 330
QY 5027 AAGGCAGGACTGAGGAGGAAGGTACAGCTGGGGGGCCCTGGGGCTTAGCTGGGACACC 5086
Db 330 ----- 330
QY 5087 CGGGGCTTCAGACAGGCGTGGCCAGGCTCTGTAGGCTTAACCTTCCTCAACACAGGA 5146
Db 330 ----- 330
QY 5147 GGAAGGAGAGTGTCCCTGGGTGCTGACCCATTGTGGGACGATGCTGTCTGCTCAGTCCGT 5206
Db 331 -----Arg--Arg 332
QY 5207 GTCCAAAGGAGATCGACGCTGATAGGCGAGGTGCGGCGACAGAGATGGGTGACCAAG 5266
Db 333 ValGlnGlnGluIleAspGluValIleGlyArgGluGlnLeuProGluMetGlyAspGln 352
QY 5267 GCTCAGATCCCTACACCACTGCGGTGATTCAGAGGTGCGAGGCTTGGGACATCATC 5326
Db 353 ThrArgMetProPheThrValAlaValIleHisGluValGlnArgPheGlyAspIleVal 372
QY 5327 CCCCTGAGGTGACCATATGATCCCTGATCGAAGTACAGGCTTCCGATCCCT 5386
Db 373 ProLeuGlyIlyValProHisMetThrSerArgAspThrGluValGlnGlyPheLeuIlePro 392
QY 5387 AAGGTAGGCTGGGGCCCTCTCACCCAGCTCAGCACGACGACCTGGTGTAGAGCCCCAG 5446
```

```
Db 393 Lys----- 393
QY 5447 CATGCTACTGCCAGTGGGGCCCACTTAGGAACCTGGCCACCTAGTCTCAATGCCAC 5506
Db 393 ----- 393
QY 5507 CACACTGACTGTCCCACTTTGGTGGGGGTTCCAGAGTATAGGCAGGGTGGCTGTCCA 5566
Db 393 ----- 393
QY 5567 TCCAGAGCCCCGCTCTAGTGGGGAGACAAACAGGACCTGCCAGAAATGTTGGAGGACCCA 5626
Db 393 ----- 393
QY 5627 GCGCTCGAGGAGAGGGGCGAGTGTGGTGCCTCTGAGAGGTGTGACTGCGCCCTGCTG 5686
Db 393 ----- 393
QY 5687 TGGGTCGAGAGGGTACTGTGAGACTTCTCGGGCGCAGGACTAGTTGACAGAGTCCAGC 5746
Db 393 ----- 393
QY 5747 TGTGTGCCAGGAGTGTGTGTCCCGGTGTGTTGGTGGCAGGGGTCCCGAGCATCCTAGA 5806
Db 393 ----- 393
QY 5807 GTCCAGTCCCCACTCTCACCTGTCATCTCTGCCAGGGAACGACACTCATCACCAACT 5866
Db 394 -----Gly-ThrThrLeuIleThrAsnIle 401
QY 5867 GTCATCGGTGTGAAGATGAGCGCTCTGGGAGAAGCCCTTCCGCTTCCACCCCGAACA 5926
Db 401 uSerSerValLeuIlySAspGluIlyValTrpIlySProPheArgPheTyrProGluIh 421
QY 5927 TTTCTGTGATGCCAGGGCCACTTTGTGAAGCGGAGGCTTCTTCCCTTCTCAGCAGG 5986
Db 421 sPheLeuAspAlaGlnGlyHisPheValIlyHisGluAlaPheMetProPhe--Ser--- 439
QY 5987 TGCTGTGGGAGCCGGCTCTCTGTCCTCCCTTCCGTTGGAGTCTTGACGGGTATCACCCA 6046
Db 439 ----- 439
QY 6047 GGAGCAGGCTCACTGACGCCCCCTCCCTCCACAGGCGCGCTGTCATGCTCGGGAG 6106
Db 440 -----AlaGlyArgArgValCysLeuGlyGlu 448
QY 6107 CCCCTGGCGCGATGAGCTTCTCTTTCACCTCCCTCTGCTGCTGAGCAGCTTTCAGCTTC 6166
Db 449 ProLeuAlaArgMetGluLeuPheLeuPheThrCysLeuLeuGlnArgPheSerPhe 468
QY 6167 TCCGTGGCGCGGAGACGCGCGCCAGCCACTCTCGTGTGCTCAGCTTCTGTTGAC 6226
Db 469 SerValProAlaGlyGlnProArgProSerAspHisGlyValPheThrPheLeuIlySVal 488
QY 6227 CCATCCCTCTACGAGCTTGTGCTGCGCCCGC 6259
Db 489 ProAlaProPheGlnLeuCysValGluProArg 499
RESULT 9
CPDK_MESAU
ID CPDK_MESAU STANDARD; PRT; 500 AA.
AC Q90Y65;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2D20 (EC 1.14.14.-) (CYP1D20).
GN CYP2D20.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
```


QY 3608 GGAGGGTGGGGCCAAAGCCCTTCATGCGAACAGCCCACTGCTCCGTCCTCCGCGCCCAAGGGGTG 3667
DB 121 -----GlyVal 122
QY 3668 ATCTCTGCTGCGCTATGCGCCCGCTGCGCGAGCAGAGCGCTTCTCCGCTGTCACCTTG 3727
DB 123 ValPheAlaArgTyrGlyProGlnTrpArgGluGlnArgPheSerValSerThrMet 142
QY 3728 CGCAACTTGGGCTGGCGCAAGAGTCTGAGACAGTGGGTGACCGAGGAGGCGCGCTGC 3787
DB 143 ArgAspPheGlyValGlyLysLysSerLeuGluGlnTrpValThrGluAlaGlyHis 162
QY 3788 CTTTGTGCGCTTCGCGGCAACAGCCGCTGGGTGATGCGCAGAGAGGCGCAAGCGGAA 3847
DB 163 LeuCysAspAla----- 166
QY 3848 CTGGGAAGCGGGGACGAGAGGCAACCCCTTACCCTGATCTCCCAACCCCAAGGAGCG 3907
DB 167 -----Phe-ThrGlnGluAlaGlyHis 173
QY 3908 CCCCTTTCCGCCCAACCGCTCTTGGACAAAGCCGTGAGCAACGTGATCGCTCCCTCAC 3967
DB 173 sProPheAsnProIleThrLeuLeuAsnLysSerValCysAsnValIleSerSerLeuIle 193
QY 3968 CTGCGGGCGCGCTTCGAGTACGACGACCTCTGCTCTCCAGCTGCTGACCTGACTCA 4027
DB 193 eTyrAlaHisArgPheAspTyrGluAspProPhePheAsnLysLeuLeuLysThrLeuGlu 213
QY 4028 GGAGGACTGMAAGGAGGTGCGGCTTCTGCGGAGGTGCGGAGCGAGAGCCGAGGAG 4087
DB 213 nGluSerPheGlyGluAspSerGlyPheIleAlaGlu----- 225
QY 4088 TCTCTGCGGGCGAGCTCTGAGAGGTGCGGGCTGCGAGCTGCGGCTCCGAAGGCGAGG 4147
DB 225 ----- 225
QY 4148 ATTGTCATAGTGGTTTGGGAAAGACATTCAGAGAGACCCACCTGTAAGAGGCGCTG 4207
DB 225 ----- 225
QY 4208 GAGGAGGGGACATCTCAGACATGTCGTGGGAGAGGTGTGCCGGGTGAGGGGGCAC 4267
DB 225 ----- 225
QY 4268 CAGGAGGCCAAGGACTCTGTACCCCGCTCCACGTTGGAGATTCGATTTTAGGTTTCT 4327
DB 225 ----- 225
QY 4328 CTTCTGGGCAAGAGAGAGGGTGGAGGCTGGGACCTTGGGAGGGGACTTGGTAGGTCACT 4387
DB 225 ----- 225
QY 4388 GGTAAGGACAGGCGCCCTGGGTCTACCTGGAGATGGCTGGGCGCTGAGACTTGTCCAG 4447
DB 225 ----- 225
QY 4448 GTGAACGCAGACACAGGAGGATTCAGACCCCGTTCTGCTGCTGAGTGTCTGAATGC 4507
DB 226 -----ValLeuAsnAl 229
QY 4508 TGTCCCGCTCTCTCGCATCCAGCGGTGGCTGGCAAGTCTACGCTTCCAAAGGC 4567
DB 229 aValProValLeuLeuArgIleProGlyLeuProGlyLysAlaPheProLysLeuThrAl 249
QY 4568 TTTCTGACCCAGCTGATGATGCTTAAGTACGACAGATGACCTGGGACCCAGCCCA 4627
DB 249 aPheMetAspSerLeuTyrLysMetLeuIleGluHisLysThrThrTrpAspProAlaGlu 269
QY 4628 GCCACCCAGACCTCTCACTGAGCCTTCTTGGCAAGAGGAGGAGGTGAGAGTGCCTGC 4687
DB 269 nProProArgGlyLeuThrAspAlaPheLeuAlaGluValGluLys----- 284
QY 4688 CACGGTGGGGGCAAGGGTGGTGGTTGAACGCTCCAGGAGGAAATGAGGGGAGGCTGGGC 4747

DB 284 ----- 284
QY 4748 AAAAGTTGGACCATGTCATCACCGCGAGCGCATCTGGGTGACAGGTGCAGAAATTG 4807
DB 284 ----- 284
QY 4808 GAGGTCAATTGGGGCTACCCCGTTCTATPCCCTGAGTATCTCTCGGCGCTGCTCAGGC 4867
DB 285 -----Al 285
QY 4868 CAAGGGAGCCCTGAGAGCAGCTTCAATGATCAGAACTTCCGATAGTGGTGAACCT 4927
DB 285 aLysGlyArgProGluSerSerPheAsnAspGluAsnLeuHisValValAlaAspLe 305
QY 4928 GTTCTTGGCGGATGTCACCACTCGACCGCTGCGCTGGGGCTCTCTGCTCATGAT 4987
DB 305 uPheIleAlaGlyMetValThrThrSerThrThrLeuSerTrpAlaLeuLeuMetIle 325
QY 4988 CCTACACCTGGATGTGACGCTGAGCCAGCTGGGGCCCAAGGCAAGGACTGAGGAGGA 5047
DB 325 eLeuHisProAspValGln----- 331
QY 5048 AGGTTACAGCTGGGGCCCTCGGCTTAGCTGGGACACCGGGGGCTTCCAGCACAGCGCT 5107
DB 331 ----- 331
QY 5108 GGCAGGCTCTGTAAGCCTAATTCTCCACACAGGAGGAGGAGAGTGTCCCTCGG 5167
DB 331 ----- 331
QY 5168 TGCTGACCCATTGTGGGAGCATGTGTCTCAGTCCGTGTCCAAACAGAGATCGACGAC 5227
DB 332 -----Ser-ArgValGlnGlnGluIleAspAsp 340
QY 5228 GTGATAGGCGCAGTGGCGGACAGAGATGGGTGACAGGCTCAATGCTTACACCACT 5287
DB 341 ValIleGlyGlnValArgArgProGluMetAlaAspGlnAlaArgMetProTyrThrAsn 360
QY 5288 GCGGTGATTCAGAGGTGCGAGCTTTGGGGACATCATCCCTGAGTGTGACCCCATATG 5347
DB 361 AlaValIleHisGluValGlnArgPheGlyAspIleAlaProValAsnValProHisMet 380
QY 5348 ACATCCCGTACATCGAAGTACAGGCTTCCGATCCCTTAAGGTAGGCTGGCGCCCTCC 5407
DB 381 ThrSerArgAspValGluValGlnGlyPheLeuIleProLys----- 394
QY 5408 TCACCCAGCTCAGCACCCAGCATCTGGTGTATAGCCCCCAGCATGGCTACTGCCAGGTGGC 5467
DB 394 ----- 394
QY 5468 CCACCTAGGAACCTGGCCACCTAGTCTCAATGCCACACACTGACTGTCCCACCTTG 5527
DB 394 ----- 394
QY 5528 GGTGGGGGTCCAGAGTATAGGAGGGCTGGCTGTGTCCATCCAGAGCCCCCTGTAGTGG 5587
DB 394 ----- 394
QY 5588 GGAGACAAACAGGACCTGCCAGAAATGTTGGAGGACCCAGCGCTGCAGGGAGAGGGGGC 5647
DB 394 ----- 394
QY 5648 AGTGTGGGTGCTCTGAGAGGTGTGACTGCGGCCCTGCTGTGGGGTCGAGAGGGTACTGT 5707
DB 394 ----- 394
QY 5708 GGAGCTTCTCGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTGCCAGGCAGTGTGTGT 5767
DB 394 ----- 394
QY 5768 CCCCCGTGTGTGTGGCAGGGGTCCAGCATCTTAGAGTCCAGTCCCACTCTCACC 5827


```
Db 1 MetGlyLeuLeuIleGlyAspAspLeuTrpAlaValValIlePheThrAlaIlePheLeu 20
QY 2129 CTCCTGGTGGACCTGATGACCGGCAACCAACGCTGGCTGCACGCTACCCGCGAGTCCC 2188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 LeuLeuValAspLeuValHisArgHisLysPheTrpThrAlaHisTyrProGlyPro 40
QY 2189 CTGCCACTCCCGGGCTGGGCAACCTTGTGCTGCATGTGGACTTCCAGAAACACACCATATG 2248
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :
Db 41 ValProLeuProGlyLeuGlyAsn-LeuLeuGlnValAspPheGluAsnMetProTyr-S 60
QY 2249 CTTCCAGCAGGTGAGGAGGAGTCTCTGGAGGGCGCGAGAGTCTCTGAGGATGCCCCACC 2308
Db ::
Db 60 er----- 60
QY 2309 ACCAGCAACATCGGTGGTGGTTAAACCAACAGGCTGGATCAGAACCCAGGCTGAGAAGG 2368
Db ----- 60
QY 2369 GGAAGCAGGTTTGGGGGACGTTCTTGGGGAGGACATTTATATCATGGCATGAAGACTGG 2428
Db ::|||
Db 61 -----LeuTyr----- 62
QY 2429 ATTTTCAAAGCCAAAGAGTAGGGCAAGGGCCTGGAGGTGGAGCTGGACTTGGCAG 2488
Db ----- 62
QY 2489 TGGGCATGCAAGCCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACA 2548
Db ----- 62
QY 2549 CCAGAAAGAAAGCCCTTGGGAATGGAAGATGATAGTCTCTGAGTGCCGTTTAAATCAGG 2608
Db ----- 62
QY 2609 AATCGAGGATGAGGGGTGAGTGCACCGGTTCAAACTTTTGCACTGTGGTCTCTG 2668
Db ----- 62
QY 2669 GGCCTCACTGCTCACCGGCATGAGCACCATCATCTGGGAATGGATGCTAACTGGGGCTCT 2728
Db ----- 62
QY 2729 CGGCAATTTGTGACTCTTGCAAGGTGATACCTGGGTGACGATCCAAACTGAGTTCCT 2788
Db ----- 62
QY 2789 CCATCACAGAAGGTGACCCCAACCCCTGCGCCACGATCAGAGGCTGGGTCTCTCTCT 2848
Db ----- 62
QY 2849 TCCACCTGCTCACTCTCTGGTAGCCCGGGGTGCTCCAAAGTTCAAATAGGACTAGGACC 2908
Db ----- 62
QY 2909 TGTAGTCTGGGTGATCTCTGGCTTGACAAGAGCCCTGACCTCCTCTGCACTTGGCG 2968
Db ::||| |||||
Db 63 -----LysLeuArgS 66
QY 2969 GCGCTTCGGGACGTGTTACGCTGACGCTGCTGGAGCGCGGTGGTCTGCTCAATG 3028
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66 erArgTyrGlyAspValPheSerLeuGlnIleAlaTrpLysProValValIleAsnG 86
QY 3029 GGTGCGCGCGTGCAGGCGGATGTCACCGCGCGGAGACAGACGCGCGGACCGCCGCG 3088
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 86 lLeuLysAlaValArgGluLeuLeuValThrTyrGlyGluAspThrAlaAspArgProL 106
QY 3089 CTGCGCCCATCTACCAAGTCTCTGGGCTTCGGGCGGTTTCCCAAGCAAGCGCGGTGG 3148
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 106 euLeuProIleTyrAsnHisLeuGlyTyrGlyAsnLysSer-Lys----- 120
QY 3149 GGACAGAGACCGGTTTCCGTGGGCGCCCGGGTGGACAGTGACCGCTAGCCCAAGACGCGCC 3208
Db ----- 120
```

```
QY 3209 GACAGGCGGTGGGGTCTTGGACGTGAAACAGAGATAAAGGCCAGCGAGTGGGCTGAGGAC 3268
Db ----- 120
QY 3269 AGTGGGCCAGGAAACCACTGCACGGGGGAGGTGCAGTCTGTGGCTGGAGGGGGCGG 3328
Db ----- 120
QY 3329 GGCTACTGCCACAGACCCGCCAGAACCCGCTGGCGAGGCTGATGCTGCAAGTGGCGGT 3388
Db ----- 120
QY 3389 GCGGGGACCGGCTATGCTGCGGCTCAGTGTGGCGGGACGGCGGGATCTTCTCTTG 3448
Db ----- 120
QY 3449 AGTGAAGAGTGTGTGAGGTGGGAGAGACGAGGTGGGCCAAACCCCGCCCGAGCAGG 3508
Db ----- 120
QY 3509 GGAGCAATGTGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGCTAGGACTG 3568
Db ----- 120
QY 3569 CGGAGACCTTGTGAGCGCCAGGTTGGAGTGGGTGGCGAGGTGGGGCCCAAGGCTT 3628
Db ----- 120
QY 3629 CATGCAAGCCCGCTGTCTCCGCCCGCCAGGGGTGATCTGTGCGGCTATGGGCC 3688
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 -----GlyValValLeuAlaProTyrGlyPro 129
QY 3689 GCGTGGCGGACGACAGCGCTTCTCCGTGTCACCTTGCACACTTGGGCTGGGCAAG 3748
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 GluTrpArgGluArgPheSerValSerThrLeuArgAspPheGlyValGlyLys 149
QY 3749 AAGTCCGTGGGACGTGGGTGACCGAGGAGCGGCTGCTTTGTGCTGCTTCCGCCGACC 3808
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 150 LysSerLeuGluGlnTrpValThrGluAlaGlyHisLeuCys-Asp----- 165
QY 3809 AAGCGGTGGGTGATGGGAGGAGCAAAAGCGGGAACCTGGGAAGCGGGGAGCGAG 3868
Db ::||| |||||
Db 166 -----ThrPheAlaLysGluAlaGlu----- 172
QY 3869 AAGGCAACCCCTTACCCGCTCTCCCGACCCCGCCAGGACCCCTTTCGCCCAACGCGCT 3928
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 173 -----HisProPheAsnProSerIleLe 180
QY 3929 CTTGACAAAGCGGTGAGCAACGTGATCGCTTCCCTCCTCCTCGGCGCGCGCTTCAGTA 3988
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 uLeuSerLysAlaValSerAsnValIleAlaSerLeuValTyrAlaArgArgPheGluTyr 200
QY 3989 CGAACACCTCTCTTCTCAGGCTCTGACCTAGCTCAGGAGGACTGAAGAGGAGTGC 4048
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 200 rGluAspProPheAsnArgMetLeuLysThrLeuLysGluSerPheGlyGluAspTh 220
QY 4049 GGGCTTCTGCGGAGGTGCGGAGCGAGACCGAGGAGTCTCTGACGCGCGAGCTCTCTG 4108
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 220 rGlyPheMetAlaGlu----- 225
QY 4109 AGAGGTGCGGGGCTGGACTGGGCTCCGAAGGCGAGGATTTGCATAGATGGGTTTGG 4168
Db ----- 225
QY 4169 AAAGACATTCAGGAGACCCCACTGTAAAGAGGCGCTGGAGGAGGAGGACATCTCAG 4228
Db ----- 225
QY 4229 ACATGTCGTGGAGAGGTGTCCCGGCTCAGGGGGCACCGAGGAGGCCAAGGACTCTG 4288
Db ----- 225
```


[1]
SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=20533996; PubMed=11083025;
RA "Okada T., Fukuhara M., Ushio F., Kurose K.;
RT "Molecular cloning and characterization of three novel cytochrome
RT P450 2D isoforms, CYP2D20, CYP2D27, and CYP2D28 in the Syrian hamster
RT (Mesocricetus auratus).";
RL Comp. Biochem. Physiol. 127C:143-152 (2000).
CC -!- FUNCTION: Has butyralol 1'-hydroxylase and debrisoquine 4-
CC hydroxylase activities.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (By
CC similarity).
CC -!- TISSUE SPECIFICITY: Expressed in liver, but not in kidney, small
CC intestine, and brain.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB031863; BAA89312.1; --
DR HSP; P00179; 1D76.
DR InterPro; IPR001128; Cytochrome P450.
DR InterPro; IPR008089; EP450_CYP2D.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01686; EP450ICYP2D.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT METAL 446 446 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 500 AA; 56489 MW; 7A5DE5878F97A954 CRC64;

Alignment Scores:
Pred. No.: 4,84e-46 Length: 500
Score: 1179.00 Matches: 346
Percent Similarity: 28.42% Conservative: 49
Best Local Similarity: 24.89% Mismatches: 92
Query Match: 6.29% Indels: 903
DB: 1 Gaps: 10

US-09-820-788A-3 (1-10278) x CPDR_MESAU (1-500)

```
QY 2468 AGGTGGAGCTGGACTTGGCAGTGGGCATGCAAGCCCATTTGGCAACATATGTTATGGAGT 2527
Db 62 ----- 62
QY 2528 ACAGAAGTCCCTTCTCTGTCACACCAAGAAAGGCTTGGGAATGGAAGATGAGTTAGTTC 2587
Db 62 ----- 62
QY 2588 CTGAGTGGCGTTTAAATCAGAAATCGAGGATGAAGGGGTGTCAGTGACCCGCTTCAAAAC 2647
Db 62 ----- 62
QY 2648 CTTTGTGCACTGTGGGTCTCTGGGCTCTCTGCTCAACCGGCATGACCATCATCTGGGAAT 2707
Db 62 ----- 62
QY 2708 GGGATGCTAACTGGGGCTCTCTGGCAATTTTGGTCACTTTTCAAGGTCAATACCTGGGTG 2767
Db 62 ----- 62
QY 2768 AGCATCCAACTGAGTTCTCTCCATCACAAGAGTGTGACCCCAACCCCTGCCCAACGAT 2827
Db 62 ----- 62
QY 2828 CAGGAGGTGGGTCTCTCTCTTCCACTGCTCTCTGTTAGTCCCGGGGGTCTGCTCAA 2887
Db 62 ----- 62
QY 2888 GGTTCAAATAGGACTAGGACCTGTAGTCTGGGGTGTCTCTGCTTGACAGAGGCCCTGA 2947
Db 62 ----- 62
QY 2948 CCTCCTCTGTCAGTGTGGGCGCCCTTTCGGGAGCTGTTTTCAGCTGTCAGCTGGCTGCA 3007
Db 63 -----LysPheArgGlnArgTyrGlyAspValPheSerLeuGlnMetAlaTrpL 79
QY 3008 CGCCGGTGTGTCTCAATGGGTGGGGCGGCGTGGCGAGCGATGTGTGACCCGCGCGC 3067
Db 79 ysProValValIleAsnGlyLeuLysAlaValArgGluValLeuValAsnCyseGly 99
QY 3068 AGGACAGCGCGACCGCGCTGCGGCTGTCAGTCTTACAGTCTCTGGGCTTGGGCGCGCTT 3127
Db 99 luAspThrAlaAspArgProValProIlePheAsnHisValGlyPheGlyHisAsnS 119
QY 3128 CCCAAGGCAAGCGGGTGGGGGACAGAGACCGCTTTCCTGGGCGCGCGGTGGACACT 3187
Db 119 erGlnGly-----ValAla-Phe----- 124
QY 3188 GACCGTAGCCCAAGCAGCGCCGACAGGCGGTGGGTCTCTGGACGTGAAACAGAGATAAG 3247
Db 124 ----- 124
QY 3248 GCCACGAGTGGGCTGAGGACAGTGGGCGCAGGAACCACTCTCACGGGGGAGGTGGGAGT 3307
Db 124 ----- 124
QY 3308 CTGTGGGCTGGAGGGGGCGGGCTACTGCCAGACCCGCCAGAAAGCCCGGTGGCGAGG 3367
Db 124 ----- 124
QY 3368 CTGATGCTCGAAGTGGCGGTGGCGGGGACCGCGCTATGTCTCGGGGTCTAGTGTGGCG 3427
Db 124 ----- 124
QY 3428 GGACGGGGGGATCTTCTTGATGGAAGAGTGTGTCAGGGTGGGCGAGAGACGAGGTGGGG 3487
Db 124 ----- 124
QY 3488 CCACACCGCGCCCGCAGGCGGGAGCAATGTGGTGTGACAAAGAGTGGGCGCTGTGCCCA 3547
Db 124 ----- 124
QY 3548 GCTGACCGGGCTAGGGACTGCGGGAGACCTTGTGGAGCGCCAGGGTGTGGAGTGGGTGCG 3607
```

Db	124	-----	124
Qy	3608	GGAGGGTGGGGCCAAAGGCTTCATGGCAACGCCCAACGTGTGCGTCCGCCCCACAGGGGTG	3667
Db	124	-----	124
Qy	3668	ATCCTGTGCGCTATGGGCCCGCGTGGCGGAGCAGAGGGCGCTTCTCCGTGTCCACCTTG	3727
Db	125	-----AlaArgTyrGlyProGlnTyrArgGluGlnArgArgPheCysValSerThrMet	142
Qy	3728	CGCAACTTGGGCGCTGGGCAAGAATCGCTGGAGCAGTGGGTGACCGAGAGAGGCCCGCTGC	3787
Db	143	ArgAspPheGlyValGlyLysLysSerLeuGluGlnTyrValThrGluGluAlaGlyHis	162
Qy	3788	CTTTGTGCGCTTCGCGCACCACCGTCGGTGGTATGGGCAGAAGGCCACAAAGCGGAA	3847
Db	163	LeuCysAspAla-----	166
Qy	3848	CTGGGAAGCGGGGACGGAGAAGCAACCCCTTACCCGCATCTCCACCCCGCAGGACG	3907
Db	167	-----Phe-ThrGlnGluAlaGlyHis	173
Qy	3908	CCCCTTTGCGCCCAACCGGCTCTTTGGACAAAGCCGTGAGCAAGTGTATCGCTCCCTCAC	3967
Db	173	sProPheAsnProThrThrLeuLeuAsnLysSerValCysAsnValIleSerLeuIle	193
Qy	3968	CTGCGGGCGCGCTTCGAGTACGACGACCCCTCGCTTCTCAGGCTGTGGACCTAGTCA	4027
Db	193	eTyrAlaHisArgPheAspTyrGluAspProPhePheAsnSerLeuLeuLysMetLeuGlu	213
Qy	4028	GGAGGGACTGAAGCAGAGAGTCCGGCTTTCTGCGCGAGGTGCGGAGCGAGACCGAGGAG	4087
Db	213	nGluSerPheGlyGluAspThrGlyPheIleAlaGlu-----	225
Qy	4088	TCTCTGCGGGCGAGCTCCTGAGAGGTGCGCGGGCTGGAATGGGGCCCTCGAAGGGCAGG	4147
Db	225	-----	225
Qy	4148	ATTGTGCATAGTGGTTTGGAAAGGACATTCACGAGAGACCCACTGTAGAAGGGCCTG	4207
Db	225	-----	225
Qy	4208	GAGGAGGAGGGACATCTCAGACATGGTCTGTGGAGAGGTGTGCCCGGTCAGGGGGCAC	4267
Db	225	-----	225
Qy	4268	CAGGAGAGGCCAAGGACTCTGTACCCCGCTCCAGCTGTGGAGATTCGATTTTAGTTTCT	4327
Db	225	-----	225
Qy	4328	CCTCTGGCAAGGAGAGAGGGGTGGAGCTGGCACTTGGGAGGGACTTGGTAGGTCAGT	4387
Db	225	-----	225
Qy	4388	GGTAAGCAGCAGGAGCGCCCTGGGTCTACTTGGAGATGGCTGGGCGCTGAGACTTGTCCAG	4447
Db	225	-----	225
Qy	4448	GTGAACGCAGACACAGGAGGATTTGAGACCCCGTTCTGTCTGGTGTAGGTGCTGAATGC	4507
Db	226	-----ValLeuAsnAl-----	229
Qy	4508	TGTCCTCGCTCTCTGCACATCCAGCGCTGGCTGGCAAGGTCCTACGCTTCCAAAGGC	4567
Db	229	aValProValLeuLeuArgIleProGlyLeuProGlyLysAlaPheProLysLeuThrAl	249
Qy	4568	TTTCTCACCCAGCTGGATGAGCTGTAACTGAGCACAGATGACCTGGGACCCAGGCCCA	4627
Db	249	aPheMetAspSerLeuTyrLysMetLeuIleGluHisLysThrThrTyrAspProAlaGlu	269
Qy	4628	GCCACCCGAGACTGACTGAGGCTTCTCTGGCAAGAAGAGGAGGTGAGAGTGGCTGC	4687

269	nProProArgGlyLeuThrAspAlaPheLeuAlaGluValGluLys-----	284
4688	CACGGTGGGGGCAAGGGTGGTGGTTGAACGTCCACAGAGGAATGAGGGGAGGCTGGGC	4747
284	-----	284
4748	AAAAGTTGGACACAGTGCATACCCGGCGAGCCGCATCTGGGTGACAGGTGCAGAAATTG	4807
284	-----	284
4808	GAGTCAITTTGGGGCTACCCGGTTCTATCCCCTGAGTATCTCTCGGCCCTCTCAGGC	4867
285	-----Al285	
4868	CAAGGGAGCCCTCAGAGCAGCTCAATGATGAGAACCTCGCATAGTGGTGGTAACTT	4927
285	alyleYargProGluSerPheAsnAepGluAsnLeuAargMetValValAlaAspMe	305
4928	GTTCCTTCGGGGATGTGACCACTCGACACCGCTGGCGTGGGGCTCTCTGCTCATGAT	4987
305	tPheIleAlaGlyMetValThrSerThrLeuSerTrpAlaLeuLeuLeuMetIl	325
4988	CCTACACCTGATGTGACGGTGTAGCCCACTGGGGCCCAAGGCAGGACTGAGGAGGA	5047
325	eLeuHisProAspValGln-----	331
5048	AGGTTACAGCTGGGGGCCCCCTGGGCTTAGCTGGGACACCCGGGCTTCCAGCACAGCGT	5107
331	-----	331
5108	GGCCAGGCTCTGTAGCCTAACTTCCTCCACACAGGAGAGAGAGTGTCCCTGGG	5167
331	-----	331
5168	TGCTGACCCATTGTGGGAGCGATGTCTGTCAGTCCGTGTCCAAAGAGAGATCGAC	5227
332	-----Ser--ArgValGlnGluIleAspAsp	340
5228	GTGATAGGGAGTGGCGCACAGAGTGGGTGACAGGCTCACATGCCCTACACCACT	5287
341	ValIleGlyValArgProGluMetAlaAspGlnAlaArgMetProTyrThrAsn	360
5288	GCCGTGATTACAGAGTGCAGCGCTTTGGGGACATCATCCCCCTGAGTGTGACCCCATATG	5347
361	AlaValIleHisGluValGlnArgPheGlyAspIleAlaProValAsnIleProHisMet	380
5348	ACATCCCGTGACATCGAAGTACAGGGCTTCCGATCCCTAAGGTAGGCTGGCGCCCTCC	5407
381	ThrSerHisAspValGluValGlnGlyPheLeuIleProllys-----	394
5408	TCACCCCACTCAGCACCAAGCACTGGTGATAGCCCCCAGCATGGCTACTGCCAGGTGGGC	5467
394	-----	394
5468	CCACTTAGGAACCTTGGCCCACTAGTCTCAATGGCCACCACTGACTGTCCCCACTTG	5527
394	-----	394
5528	GGTGGGGGTCCAGAGTATAGGCAGGGCTGGCTGTCCATCCAGAGCCCGCTCTAGTGG	5587
394	-----	394
5588	GGAGACAAACAGGACCTGCCAGNATGTTGGAGGACCCAGCGCTCGAGGAGAGGGGGC	5647
394	-----	394
5648	AGTGTGGGTGCTCTGAGAGGTGTGACTGCGCCCTGCTGTGGGGTCCGAGAGGGTACTGT	5707
394	-----	394
5708	GGAGCTTCTCGGGCGCAGAGCTAGTTGACAGATCCAGCTGTGTGTGCCAGCGATGTGTGT	5767
394	-----	394

Db 61 ----- 61
QY 2453 AGGCAAGGCGCTGGAGTGGAGCTGGACTTGGCAGTGGGCATGCAAGCCCATGGGCAA 2512
Db 61 ----- 61
QY 2513 CATATGTTATGAGTACAAAGTCCCTTCTGTGTGACACCAAGAAAGCCCTTGGGAATG 2572
Db 62 ----- 63
QY 2573 GAAGATGAGTTAGTCTGAGTGCCTTTAAATCACAAATCGAGATGAAGGGGTGCAG 2632
Db 63 ----- 63
QY 2633 TGACCGGTTCAACCTTTTGCATGTGTGGTCTCTGGGCTCCTGCTCAACCGGCATGGA 2692
Db 63 ----- 63
QY 2693 CCATCATCTGGGAATGGGATGCTAACTGGGCTCTCGCAATTTTGGTACTCTTGCAA 2752
Db 63 ----- 63
QY 2753 GGTACATCTGGGTGACGCATCCAACTGAGTTCTCCATCACAGAAGTGTGACCCCA 2812
Db 63 ----- 63
QY 2813 CCCCTGCCCCAGCATCAGAGGCTGGTCTCTCTCTCCACTGCTCACTCTGTAGCC 2872
Db 63 ----- 63
QY 2873 CCGGGGTGTTCCAAAGTTCAATAGGACTAGGACTGTAGTCTGGGTGATCTGGCTT 2932
Db 63 ----- 63
QY 2933 CACAAGAGGCGCTGACCCCTCTCTGAGTTGCGGCGCGCTTCGGGACGTTTCAGCC 2992
Db 64 ----- 74
QY 2993 TGCAGTGGCTGGAGCGCGGTGTGCTCAATGGCTGGCGCGCTGCGCGAGGCGGA 3052
Db 74 euGlnMetGlyTriplysProValValIleAsnGlyLeuLysAlaValGlnGluLeuL 94
QY 3053 TGGTACCGCGGCGGACGACGCGCGCGCTGCTCAATGGCTGGCGCGCTGCGCGAGGCTGG 3112
Db 94 euValThrCysGlyGluAspThrAlaAspArgProGluMetProIlePheGlnHisIleG 114
QY 3113 GCTTCGGGCGCGTTTCCCAAGGCAAGCGCGGTGGGGGACAGACCGCGTTTTCGTGGG 3172
Db 114 lyTyGlyHisLysAla-Lys----- 120
QY 3173 CCGCGGTGGAGTGAAGTGGTCCAGCCAGACGCGCGAGCGGCTGGGCTCTGGAGT 3232
Db 120 ----- 120
QY 3233 GAAACAGAGATAAGGCCAGCGAGTGGGCTGAGGACAGTGGGCCAGGAAACCACTGCAC 3292
Db 120 ----- 120
QY 3293 GGGGAGGTGCGAGTCTGTGGCTGGGAGGGGGGGGCTACTGCGCCAGACCCGCGCAA 3352
Db 120 ----- 120
QY 3353 GCCCGTGGCGAGGCTGATGCTGCGAAGTGGCGGTGGCGGGACCGCGCTATGCTGCG 3412
Db 120 ----- 120
QY 3413 GGCTCAGTGTGGCGGGACGGGCGGATCTTCTTGTAGTGAAGGTGGTCAGGGTGGGC 3472
Db 120 ----- 120
QY 3473 AGAGACGAGGTGGGCGCAAAACCCCGCCAGCGAGGGGAGCAATGTGGGTGAGCAAGAG 3532

Db 120 ----- 120
QY 3533 TGGGCGCTGTGCCCAAGCTGGACCGGGCTAGGGACTGCGGGAGACCTTGTGGAGCCCAAG 3592
Db 120 ----- 120
QY 3593 GTTGGAGTGGGTGGCGAGGGTGGGCCAAGCCCTTATGGCAACGCCACGCTGTCCGTC 3652
Db 120 ----- 120
QY 3653 CCGCCCCCAGGGGTGATCTGTGCGCTATGGGCGCGCGTGGCGAGCAGAGCGCTTC 3712
Db 121 ----- 137
QY 3713 TCCGTGTCCACTTGGCGCACTTGGGCTGGCGCAAGAGTGCCTGGAGCAGTGGGTGACC 3772
Db 138 SerValSerThrLeuArgAsnPheGlyValGlyLysSerLeuGluGlnTrpValThr 157
QY 3773 GAGGAGCGCGCTGCTTGTGCGCTTCGCGCCAGCCGCTGGGTGGGTGAGGCGAGAAG 3832
Db 158 AspGluAlaSerHisLeuCysAspAla----- 166
QY 3833 GGCACAAAGCGGAACTGGGAAGCGCGGGACGGAGAGGCAACCCCTTACCCGATCTC 3892
Db 167 --LeuThrAlaGluAlaGlyArg----- 173
QY 3893 CCCACCCCGAGGACGCGCTTTCGCCCAACGCGCTCTTGGACAAAGCCGTGAGCAAGCT 3952
Db 174 -----ProLeuAspProTyThr-LeuLeuAsnLysAlaValCysAsnVa 188
QY 3953 GATCGCTCCCTCACCTCGCGGCGCGCTTCGAGTACGACGACCTCGCTTCTCCTCAGCT 4012
Db 188 lileAlaSerLeuileTyralaArgPheAspTyrglyAspProaspPheIleLysVa 208
QY 4013 CTTGACCTAGCTCAGGAGGACTCAAGAGGAGTTCGGGCTTTCGCGGAGTGGCGAG 4072
Db 208 lLeuLysIleLeuLysGluSerMetGlyGluGlnThrGlyLeuPheProGlu----- 225
QY 4073 CGAGAGACCGAGGAGTCTCTCGAGGGGAGCTCTCGAGAGGTGCGGGGCTGGAGTGGGG 4132
Db 225 ----- 225
QY 4133 CCTCGAAGGCGAGGATTTGCATAGATGGGTTTGGAAAGGACATTCAGGAGACCCAC 4192
Db 225 ----- 225
QY 4193 TGTAAGAAGGCGCTGGAGGAGGGGACATCTCAGACATGCTGCTGGGAGAGGTGTGCC 4252
Db 225 ----- 225
QY 4253 CGGCTCAGGGGGCACCCAGGAGAGGCCAAGGACTCTGTACCCCGCTCCACGTGGAGATT 4312
Db 225 ----- 225
QY 4313 CGATTTTAGGTTTCTCTCTGGGCAAGGAGAGGGGTGGAGCTGGCACTTTGGGAGGGA 4372
Db 225 ----- 225
QY 4373 CTTGGTGAGGTGAGTGGTAAGGACAGGCGAGGCCCTGGGTCTACCTGAGATGGCTGGGC 4432
Db 225 ----- 225
QY 4433 CTGAGACTTGTCCAGGTGAACGCAGAGCAAGGAGGATTTGAGACCCCGTTCTGTCTGGT 4492
Db 225 ----- 225
QY 4493 GTAGGTGCTGAATGCTGTCTCCCGTCTCTCTGACATCCAGCGCTGGTGGCAAGTCTCT 4552
Db 226 ----ValLeuAsnMetPheProValLeuLeuArgIleProGlyLeuAlaAspLysValPh 244
QY 4553 ACGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGTGTCTAACTGAGCAGAGATGAC 4612
Db 244 eProGlyGlnLysThrPheLeuThrMetValAspAsnLeuValThrGluHisLysLysTh 264

```

4613 QY CTGGGACCCAGCCACCCGAGACCTGACTGAGGCTTCTGCGCAAGAGAGGAA 4672
      Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
264 QY rtrpAspProAspGlnProAspGlnProAspGlnProAspGlnProAspGlnPro 284
      Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4673 QY GGTGAGAGTGGTCCACCGTGGGGGCAAGGGTGGTGGTGAACGTCCTCCAGGGAAT 4732
      Db |
284 QY |
      Db |
4733 QY GAGGGAGGCTGGGCAAAAGTTGGACAGTGCATCACCCGGGAGCCGATCTGGGCTG 4792
      Db |
284 QY |
      Db |
4793 QY ACAGGTGCAGAAATTGAGGTCAATTGGGGCTACCCCGTTCTATCCCTGAGTATCCTCT 4852
      Db |
284 QY |
      Db |
4853 QY CGGCTCTGCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCATGATGAGAACTCGGCGAT 4912
      Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
285 QY -----AlaLysGlyAsnProGluSerPheAsnAspAlaAsnLeuArgLue 300
      QY AGTGGTGGTAACTCTCTCTCGGGGATGGTACCCTCGACACAGCTGCGCTGGGG 4972
      Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 QY uValValAsnAspLeuPheLysAlaGlyMetValThrSerIleThrLeuThrTrpAl 320
      QY CCTCTGCTCATGATCCTACACCTGATGTGCAGCGTGCAGCCAGCTGGGGGCCAAGGCA 5032
      Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
320 QY aLeuLeuLeuMetIleLeuHisProAspValGln----- 331
      QY GGGACTGAGGGAGGAAGGGTACAGCTGGGGGCCCCCTGGGCTAGCTGGGACACCCGGGGC 5092
      Db |
331 QY |
      Db |
5093 QY TTCCAGCACAGGCGTGGCCAGGCTCTGTAAAGCTTAATCTCTCCAACACAGAGGAAGG 5152
      Db |
331 QY |
      Db |
5153 QY AGAGTGTCCCTGGGTGCTGACCCATTTGTGGGAGCGCATGTCTGTCCAGTCCGTGTCCAA 5212
      Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
332 QY -----Cys--ArgValGln 335
      QY CAGGAGATCAGACCTGATAGGCGCAGGTGGCGGACACAGATGGGTGACAGGCTCAC 5272
      Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
336 QY GlnGluIleAspGluValIleGlyGlnValArgHisProGluMetAlaAspGlnAlaHis 355
      QY ATGCCCTACCACTGCGCTGATTCACAGGTGCAGCGCTTTGGGGACATCATCCCGCTG 5332
      Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
356 QY MetProPheThrAsnAlaValIleHisGluValGlnArgPheAlaAspIleValProMet 375
      QY AGTGTGACCCATATGATACCTCCGTGACATCGAAGTACAGGGCTTCGCGATCCTTAAGGTA 5392
      Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
376 QY AsnLeuProHisLysThrSerArgAspIleGluValGlnGlyPheLeuIleProLys--- 394
      QY GGCCTGGCGCCCTCTCTCACCCAGCTCAGCACACGACCTGGTGTAGTCCCGCCAGCATGSC 5452
      Db |
394 QY |
      Db |
5453 QY TACTGCCAGGTGGGCCACTCTAGGNAACCTGGCCACCTAGTCTCAATGCCACCACT 5512
      Db |
394 QY |
      Db |
5513 QY GACTGTCCCACTTGGTGGGGGTCCAGATATAGGCAGGCTGGCGTGTCCATCCAGA 5572
      Db |
394 QY |
      Db |
5573 QY GCCCGCTCTAGTGGGGAGACAAACAGGACCTGCCAGAAATGTTGGAGGACCCAGCGCT 5632
      Db |
394 QY |
      Db |
5633 QY GCAGGAGAGGGGCGAGTGGGTGCTCTGAGAGGTGTGACTGGCGCTGCTGTGGGT 5692
      Db |
394 QY |
      Db |
```

5693 CGGAGAGGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTG 5752

394 ----- 394

5753 CCAGGACAGTGTGTGTCCTCCCGCTGTGTGGTGGCAGGGGTCCAGAGCATCTAGAGTCCAG 5812

394 ----- 394

5813 TCCCACTCTCACCTGCATCTCTGTCGCCAGGGAACGACACTCATCACCAACCTGTATC 5872

395 -----Gly-ThrThrLeuIleProAsnLeuSerSe 404

5873 GGTGCTGAAGATGAGCGCGTCTGGAGAGCGCTTCCGCTTCCACCCGAAACACTTCTCT 5932

404 rValLeuLysAspGluThrValTrpGluLysProLeuArgPheHisProGluHisPheLe 424

5933 GATGCCCGAGGCGCACTTTGTGAAGCGGGAGCGCTTCTGCTTCTCTGACGAGGTCCCTG 5992

424 uAspAlaGlnGlyAsnPheValLysHisGluAlaPheMetProPhe--Ser----- 440

5993 TGGGAGCGCCGCTCCCTGTCCTCTCGTGGAGTCTTGCAGGGGTATCACCAGGAGCC 6052

440 ----- 440

6053 AGGCTCACTGACGCGCCCTCCCTCCCAAGGCGCGCGTGCATGCTCGGGAGCGCCCTG 6112

441 -----AlaGlyArgAlaCysLeuGlyGluProLeu 451

6113 GCGCCATGAGAGCTTCTCTTCTTACCTCCCTGCTGTCGACGACTTTCAGCTTCTCCGTG 6172

452 AlaArgMetGluLeuPheLeuPheThrCysLeuLeuGlnArgPheSerPheSerVal 471

6173 GCGCGCGGACGCGCGCGCGCGCGCTCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6232

472 ProThrGlyGlnProArgProSerAspTyrGlyValPheAlaPheLeuLeuSerProSer 491

6233 CCTACGAGCTTGTGTGCTGCGCGCGC 6259

492 ProTyrGlnLeuCysAlaPheLysArg 500

RESULT 13

CPDG_CAVPO STANDARD; PRT; 500 AA.

AC Q64403; OS4866;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cytochrome P450 2D16 (EC 1.14.14.1) (CYP2D16).

GN CYP2D16.

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.

OX NCBI_TaxID=10141;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-38.

RC STRAIN=13; TISSUE=Adrenal cortex;

RX MEDLINE=95251703; PubMed=7733969;

RA Jiang Q., Voigt J.M., Colby H.D.;

RT "Molecular cloning and sequencing of a guinea pig cytochrome P4502D (CYP2D16): high level expression in adrenal microsomes.";

RL Biochem. Biophys. Res. Commun. 209:1149-1156(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Hartley white; TISSUE=Adrenal gland;

RA Sun Y., Voigt J.M., Pierce J.C., Colby H.D.;

RT "The gene sequence of a xenobiotic metabolism-related cytochrome P450 isozyme (CYP2D16) in guinea pig adrenal gland."

RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate monooxygenases. In liver microsomes, this enzyme is involved in an NADPH-dependent electron transport pathway. It oxidizes a variety of structurally unrelated compounds, including steroids, fatty acids, and xenobiotics.

CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE INNER ZONE OF
CC THE ADRENAL CORTEX.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U21486; AA68479.1; -.
CC DR EMBL; AF020345; AAB94568.1; -.
CC DR PIR; JC4153; JC4153.
CC DR HSSP; P00179; 1DT6.
CC DR InterPro; IPR001128; Cytochrome P450.
CC DR InterPro; IPR008069; EP450_CYP2D.
CC DR Pfam; PF00067; P450; 1.
CC DR PRINTS; PR01686; EP450ICYP2D.
CC DR PRINTS; PR00385; P450.
CC DR PROSITE; PS00086; CYTOCHROME P450; 1.
CC KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
CC Microsome; Endoplasmic reticulum.
CC FT DOMAIN 81 84 POLY-VAL.
CC FT METAL 446 446 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC FT CONFLICT 123 123 I -> V (IN REF. 2).
CC FT CONFLICT 127 127 Y -> N (IN REF. 2).
CC FT CONFLICT 148 148 G -> R (IN REF. 2).
CC SQ SEQUENCE 500 AA; 55800 MW; 2429247E49BF6B24 CRC64;

Alignment Scores:

Pred. No.:	2,95e-45	Length:	500
Score:	1161.50	Matches:	350
Percent Similarity:	28.44%	Conservative:	49
Best Local Similarity:	24.95%	Mismatches:	98
Query Match:	6.20%	Indels:	906
DB:	1	Gaps:	12

US-09-820-788a-3 (1-10278) x CPDG_CAVPO (1-500)

QY	2078	ATGGGGCTA-----GAAGCATGTGGCCCTGGCCATGATAGTGGCATCTTCTCTG	2128
DB	1	MetGlyLeuLeuThrGlyAspAlaLeuPheSerValAlaValAlaValAlaLeuPheLeu	20
QY	2129	CTCCTGGTGGACCTGATGACCGGCACCAACGCTGGCTGCACGCTACCCGCCAGGTCCC	2188
DB	21	LeuLeuValAspLeuMetHisArgArgGlnArgTrpAlaAlaArgTrpProProGlyPro	40
QY	2189	CTGCCACTGCCGGCTGGCAACCTCTGTCATGTGGACTTCCAGAACACACCATACTG	2248
DB	41	ValProValProGlyLeuGlyAsn-LeuLeuGlnValAspPheGluAsnMetAlaTySe	60
QY	2249	CTTGACACAGGTGAGGAGGAGGTCTCGAGGGGGGAGAGGTCTGAGGATGCCCCACC	2308
DB	60	r-----	60
QY	2309	ACCAGAAACATGGTGGTGTAAACACAGGCTGGATCAGAGCCAGGCTGAGAAGG	2368
DB	60	-----	60
QY	2369	GGAAGCAGGTTTGGGGACGTTCTCTGGGGAAGGACATTATACATGGCATGAAGACTGG	2428
DB	60	-----	60
QY	2429	ATTTTCCAAAGCCAAAGAGTAGTAGGCAAGGGCCTGGAGGTGAGCTGGACTTGGCAG	2488
DB	60	-----	60
QY	2489	TGGGCATCAAGCCCATTTGGGCAACATATGTTATGAGTACAAAGTCCCTTCTGCTGACA	2548

DB	61	-----	11111111	Cys-Asp-	62
QY	2549	CCAGAAGAAAGGCGCTTGGGAATGGAAGATGAGTTAGTCTCTGAGTGGCGTTTAAATCAGC	2608		
DB	62	-----	-----	-----	62
QY	2609	AAATCGAGGATGAAGGGGTGTCAGTACCCCGTTCAAACCTTTTGCACTGTGGGTCTCTCG	2668		
DB	62	-----	-----	-----	62
QY	2669	GGCCTCACTGCTACCGGCATGAGCACATCATCTGGGAATGGATGCTAACTGGGGCCTCT	2728		
DB	62	-----	-----	-----	62
QY	2729	CGGCAATTTTGGTGACTCTTGTCAAGGTATACCTGGGGTGACGCATCCAAACTGAGTTCT	2788		
DB	62	-----	-----	-----	62
QY	2789	CCATCACAAGAGTGTGACCCCAACCCCTGCCCAAGATCAGGAGGTGGGTCTCTCTCT	2848		
DB	62	-----	-----	-----	62
QY	2849	TCCACCTGCTCACTCTCTGTAGCCCGGGGTGCTCCAAGGTTCAAATAGGACTAGGACC	2908		
DB	62	-----	-----	-----	62
QY	2909	TGTAGTCTGGGGTGATCTTGTGCAAGAGGCCCTGACCCCTCTCTGAGTTGGCGC	2968		
DB	63	-----	-----	-----	66
QY	2969	CCCGCTTGGGGACGTCTTCACTGCTGAGCTGGCGCTGGACCGCCGGTGGTCTGCTCAATG	3028		
DB	66	isGlnPheGlyAspValPheSerLeuGlnPheValTrpThrProValValValAsnG	86		
QY	3029	GGCTGGGGCGCTGCGGAGCGCATGGTGACCCGCGGCGAGACACGCGCCACCGCCGC	3088		
DB	86	lyLeuLeuAlaValArgGluAlaLeuValAsnAsnSerThrAspThrSerAspArgProT	106		
QY	3089	CTGCGCCCATCTACAGGTCTCTGGGCTTGGCGCTGCGCTGCCAGAGCAAGCGGGTGGG	3148		
DB	106	hrLeuProThrAsnAlaLeuLeuGlyPheGlyProLysAlaGlnGlyValIle-Gly---	124		
QY	3149	GGACAGAGACCGCGTTTCCGTGGGGCCCGGTGACAGTACCTGAGCCCAAGCAGCGCC	3208		
DB	124	-----	-----	-----	124
QY	3209	GACAGGCGTGGGGTCTTGGACGTGAAACAGAGATAAGGCCAGCGAGTGGGCTGAGGAC	3268		
DB	124	-----	-----	-----	124
QY	3269	AGTGGCGAGAAACACCTGTCACGGGGAGGTGCGAGTCTGTGGGTGGAGGGGGCGG	3328		
DB	124	-----	-----	-----	124
QY	3329	GGCTACTGCCACAGACCCGCCAAGAACCCGGTGGGGGAGGCTGATCGTCGAAGTGGCGGT	3388		
DB	124	-----	-----	-----	124
QY	3389	GGCGGGACCGCGCTATGCTGCGGGCTAGTGTGGCGGACGCGGGGATCTTCTCTTG	3448		
DB	125	-----AlaTyr-----	-----	-----	126
QY	3449	AGTGGAAAGTGTGTCAGGGTGGCAGAGACGAGGTGGGGCCAAACCCCGCCAGGACAGG	3508		
DB	126	-----	-----	-----	126
QY	3509	GGACATGTGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGGCTAGGGACTG	3568		
DB	126	-----	-----	-----	126
QY	3569	CGGAGAGACCTTGTGGAGCGCCAGGTTTGGAGTGGGTGGCGAGGGTGGGGCCAAAGGCCTT	3628		

[illegible]

QY	5789	GGGTCCAGCATCTGATAGTTCAGTCCCACCTCTCACCTGCATCTCCTGCCAGGGAAC	5849
Db	395	-----	396
			Gly-Th
QY	5849	GACACTCATCAACAACCTGTCTATCGTGTCGAAAGGATGAGCCGTCTGGGAGAGCCCTT	5908
Db	396	rMetLeuPheThrAsnLeuSerValLeuLysAspGluThrValTrpGlnLysProLe	416
QY	5909	CCGCTTCCACCACCGAACAACCTTCTCTGGATGCCAGGCGCACCTTTGTGAAGCCGGAGGCCTT	5968
Db	416	uHisPheHisProGlyHisPheLeuAspAlaGluArgPheValLysArgGluAlaPh	436
QY	5969	CCTGCGCTTTCTACGAGGTGCCTGTGTGGGAGACCGCGCTCCCTGTCCCTTCCGTGGAGTC	6028
Db	436	eMetProPheSerAlagly	442
QY	6029	TTCAGGGGTATCACCCAGGAGCCAGGCTCACTGACGCCCCCTCCCTCCCCACAGGCCGC	6088
Db	443	-----	443
		-----	Pro--
QY	6089	CGTGCATGCTCGGGAGACCCCTGGCCCGCATGGAGCTCTTCTCTCTTTCACTCCCTG	6148
Db	444	ArgIleCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu	463
QY	6149	CTCAGACATTCACTGCTTCTCGTGGCGCCGCGACAGCCCGCGCACCACTCTCTGTGTC	6208
Db	464	LeuGlnArgPheSerPheSerValProGluGlyGlnProArgProSerAspArgGlyAla	483
QY	6209	GTCACTTCTGTGACCCCATCCCCCTACGAGCTTGTGCTGTGCCCCGC	6259
Db	484	ProTyrlauValValLeuProSerProTyrGlnLeuCysAlaValLeuArg	500
		::::	
RESULT 14			
CPD5_RAT		STANDARD;	PRT; 504 AA.
ID	CPD5_RAT		
AC	P12939;		
DT	01-OCT-1989 (Rel. 12, Created)		
DT	01-OCT-1989 (Rel. 12, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Cytochrome P450 2D5 (EC 1.14.14.1) (CYPIID5) (P450-DB5) (P450-CMF1B)		
DE	(Debrisoquine 4-hydroxylase).		
GN	CYP2D5 OR CYP2D-5.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.		
NCBI_TaxID=10116;			
[1]			
SEQUENCE FROM N.A.			
STRAIN=Sprague-Dawley; TISSUE=Liver;			
MEDLINE=90189185; PubMed=2107330;			
RA Matsuura T., Zanger U.M., Hardwick J.P., Gelboin H.V., Meyer U.A.,			
Gonzalez F.J.;			
"The rat P450 IID subfamily: complete sequences of four closely			
linked genes and evidence that gene conversions maintained sequence			
homogeneity at the heme-binding region of the cytochrome P450 active			
site.";			
J. Mol. Evol. 30:155-169(1990).			
[2]			
SEQUENCE FROM N.A.			
MEDLINE=90057430; PubMed=2819073;			
RA Matsunaga T., Zanger U.M., Hardwick J.P., Gelboin H.V., Meyer U.A.,			
Gonzalez F.J.;			
"The CYP2D gene subfamily: analysis of the molecular basis of the			
debrisoquine 4-hydroxylase deficiency in DA rats.";			
Biochemistry 28:7349-7355(1989).			
[3]			
SEQUENCE FROM N.A.			
MEDLINE=89366685; PubMed=2771656;			
RA Ishida N., Inuzuka C., Tawaragi Y., Sugita O., Nakazato H.,			
Noguchi T., Sassa S., Kappas A.;			
"Cytochrome P450CWF cDNA: nucleotide sequence of P450CWF1b.";			
Nucleic Acids Res. 17:6407-6407(1989).			
[4]			
SEQUENCE OF 18-504 FROM N.A.			
RP			

```
Db 61 ----- 61
QY 2393 TGGGGAAGGACATTTTATACATGSCATGAAGGACTGGATTTCACAAAGGCCAAGGAAGT 2452
Db 61 ----- 61
QY 2453 AGGGCAAGGCCCTGGAGGTGGAGCTGGACTTGGCAGTGGGCATGCAAGCCCATTTGGGCAA 2512
Db 61 ----- 61
QY 2513 CATATCTATGGAGTACAAAGTCCCTTCTGCTGACACACAGAAAGGCCCTTGGGAATG 2572
Db 62 ----- 63
QY 2573 GAAGATGAGTTACTCTGAGTGCCTTTAAATCACGAAATCGAGGATGAAGGGGTGCAG 2632
Db 63 ----- 63
QY 2633 TGACCCGGTTCAAACCTTTTGCACTGTGGTCTCTCGGGCTCACTGCTCACCGGCATGA 2692
Db 63 ----- 63
QY 2693 CCATCATCTGGGAATGGATGCTAACTGGGGCTCTCGGCAATTTTGTGACTCTTGCAA 2752
Db 63 ----- 63
QY 2753 GGTCTATACCTGGGTGACGATCCAAACTGAGTTCTCTCCATCACAGAAGGTGTGACCCCA 2812
Db 63 ----- 63
QY 2813 CCCCTGCCACGATCAGGAGGTGGGTCTCTCTTCCACTGCTCACTCTGTPAGCC 2872
Db 63 ----- 63
QY 2873 CCGGGGTCTGTCAGGTTCAATAGACTAGGACCTGTAGTCTGGGGTATCCTGGCTT 2932
Db 63 ----- 63
QY 2933 GACAGAGGCCCTGACCTCCCTCTGCAGTTGGGGCGCTTCGGGGACGTGTTTCAGCC 2992
Db 64 ----- 74
QY 2993 TGCAGTGGCTGACGCGGTGGTGTCTCAATGGGTGGCGCGCTTCGCGAGGCGA 3052
Db 74 euGlnMetGlytPlysProMetValIleValAenArgLeuLysAlaValGlnGluVal 94
QY 3053 TGGTGACCGCGGAGGACAGCGCCGACCGCGCGCTGGCCCATCTACAGGTCTCTGG 3112
Db 94 euValThrHisGlyGluAspThrAlaAspArgProValProIlePheLysCysLeu 114
QY 3113 GCTTCGGCGCGCTTCCCAAGGCAAGCGCGGTGGGGGACAGAGACCGCGTTTCCGTGG 3172
Db 114 lyValLysProArgSer-Gln----- 120
QY 3173 CCCCCGGTGGAGTGAACCTAGCCCAAGCGCGCGAGCGGCGTGGGGTCTTGGAAGT 3232
Db 120 ----- 120
QY 3233 GAAACAGAGATAAAGCCAGCGAGTGGGTGAGGACAGTGGGCCAGGAAACCACTGCAC 3292
Db 120 ----- 120
QY 3293 GGGGAGGTGCGAGTCTGTGGGTGGAGGGGGCGGGCTACTGCCCCAGACCGCCAGAA 3352
Db 120 ----- 120
QY 3353 GCCCGTGGGCGAGGCTGATGCGTCGAAGTGGCGGTGGGGGACCGGCCTATGCTGCG 3412
Db 120 ----- 120
QY 3413 GGCTCAGTGTGGCGGACGGGGGGGATCTTCTTGAAGTGGAAAGGTGTGAGGGTGGGC 3472
Db 120 ----- 120

QY 3473 AGAGACGAGGTGGGGCCAAACCCCGCCAGCGAGGGAGCAATGTGGGTGAGCAAGAG 3532
Db 120 ----- 120
QY 3533 TGGGCCCTGTGCCAGCTGGACCGGGCTAGGACTCTGGGAGACCTTTGTGGAGCGCCAGG 3592
Db 120 ----- 120
QY 3593 GTTGAGTGGGTGGCGAGGGTGGGGCCAAAGCCCTTCATGGCAACGCCACGTCGTCCGTC 3652
Db 120 ----- 120
QY 3653 CCGCCCCCAGGGGTGATCCTGTGCGCTATGGGCCCGGTGGCGAGCAGAGCGCTTC 3712
Db 121 ----- 137
QY 3713 TCGGTGTCCACCTTCGGCAACTTGGGCTGGGCAAGAGTCTCTGGAGCAGTGGGTGACC 3772
Db 138 SerValSerThrLeuArgThrPheGlyMetGlyLysLysSerLeuGluGluTrpValThr 157
QY 3773 GAGGAGCGCGCTGCTTGTGCCCTTCGCCGACCAAGCCGCTGGGTGATGGGCAAG 3832
Db 158 lysGluAlaGlyHisLeuCys-----Asp-AlaPheTh 168
QY 3833 GGCACAAACCGGAACTGGGAAGCGGGGACGAGAGGCAACCCCTTACCCGCTATCTC 3892
Db 168 rAlaGlnAen----- 171
QY 3893 CCCACCCCGAGACGCCCTTTCGCCCCCAACCGCTCTTGGACAAAGCCGTGAGCAACGT 3952
Db 172 ----- 188
QY 3953 GATCGCTCCCTCACCTGCGGGCGCTTCGAGTACGACGACCTCTCGCTTCTCAGGCT 4012
Db 188 lIleAlaSerLeuIlePheAlaArgPheGluArgPheGluAspProTyLeuIleArgme 208
QY 4013 GCTGACCTTAGCTCAGAGGAGTGAAGGAGAGTCTGGGCTTCTGCGGAGGTGGGAG 4072
Db 208 tLeuThrLeuValGluGluSerLeuIleGluValSerGlyPheIleProGlu----- 225
QY 4073 CGAGAGACGAGGAGTCTCTGCAGGGCGAGCTCTGAGAGGTGCGGGGCTGAGCTGGGG 4132
Db 225 ----- 225
QY 4133 CCTCCGAAGGCGAGGATTTCATAGTGGTTTGGGAAGGACATTCCAGGAGACCCAC 4192
Db 225 ----- 225
QY 4193 TGTAAGAAGGGCTGAGAGGAGGGGACATCTCAGACATGCTGCTGGGAGAGGTGTGCC 4252
Db 225 ----- 225
QY 4253 CCGGTGAGGGGACACAGAGAGGCCAAGGACTCTGTACCCCGCTCCACGTTGGAGATT 4312
Db 225 ----- 225
QY 4313 CGATTTTAGTTTCTCTCTGGGCAAGGAGAGGGGTGGAGGCTGCACTTGGGGAGGGA 4372
Db 225 ----- 225
QY 4373 CTTGGTGGGTCAAGTAAAGCAGGAGCGGCTGGGTCTTACCTGGAGATGGCTGGGGC 4432
Db 225 ----- 225
QY 4433 CTGAGACTTGTCCAGGTGAACGACAGAGCACAGAGGGATTTGAGACCCCGCTTCTGTCTGGT 4492
Db 225 ----- 225
QY 4493 GTAGTGTGTGATGTGTCTCCCGTCTCTCTGACATCCAGCGCTGGCTGGCAAGGTCTCT 4552
Db 226 -----ValLeuAenThrPheProAlaLeuLeuArgIleProGlyLeuAlaAspLysValPh 244
```


Biochemistry 28:4779-4784(1989).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=89123394; PubMed=2914938;
 RA Wong G., Itakura T., Kawajiri K., Skow L., Negishi M.,
 RT "Gene family of male-specific testosterone 16 alpha-hydroxylase (C-P-
 RT 450(16 alpha)) in mice. Organization, differential regulation, and
 RT chromosome localization.";
 RL J. Biol. Chem. 264:2920-2927(1989).
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
 CC monooxygenases. In liver microsomes, this enzyme is involved in an
 CC NADPH-dependent electron transport pathway. It oxidizes a variety
 CC of structurally unrelated compounds, including steroids, fatty
 CC acids, and xenobiotics.
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -!- INDUCTION: P450 can be induced to high levels in liver and other
 CC tissues by various foreign compounds, including drugs, pesticides,
 CC and carcinogens.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M23998; AAA40427.1; -;
 DR EMBL; M23997; AAA40428.1; ALT_SEQ.
 DR EMBL; M27168; AAA39876.1; -;
 DR EMBL; M24267; AAA75462.1; -;
 DR EMBL; M24262; AAA75462.1; JOINED.
 DR FIR; B27384; B27384.
 DR FIR; S15806; A27384.
 DR HSSP; P00179; 1DT6.
 DR MGD; MGI:88606; Cyp2d9.
 DR InterPro; IPR001128; Cytochrome P450.
 DR InterPro; IPR008069; EP450_CYP2D.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01686; EP450ICYP2D.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT METAL 446 446 IRON (HEME AXIAL LIGAND).
 FT CONFLICT 54 55 LG -> OD (IN REF. 2).
 SQ SEQUENCE 504 AA; 56949 MW; B85862205472A176 CRC64;
 Alignment Scores:
 Pred. No.: 3.7e-44 Length: 504
 Score: 1137.00 Matches: 337
 Percent Similarity: 28.83% Conservative: 66
 Best Local Similarity: 24.11% Mismatches: 92
 Query Match: 6.07% Indels: 903
 DB: 1 Gaps: 10
 US-09-820-788a-3 (1-10278) x CPD9_MOUSE (1-504)
 QY 2093 CTGGTCCCTGGCCATAGTAGGCGCATCTTCTGCTGCTGGTGGACCTGATGCACCGG 2152
 DB 9 LeuTrpProValAlaIlePheThrValIlePheIleLeuLeuValAspLeuThrHisGln 28
 QY 2153 CACCAACGTGGCTGCACGCTACCGCAGGTGCCCTGCCACTGCCGGGTGGGCAAC 2212
 DB 29 ArgGlnArgTrpThrSerArgTyrProGlyProValProTrpProValLeuGlyAsn 48
 QY 2213 CTGCTGCATGTGGACTTCAGACACACATCTGCTCGACCGGTGAGGAGGAGGT 2272
 DB 49 -LeuLeuGlnValAspLeuGlyAsnMetProTyr-Ser----- 60

2273 CCTGGAGGGCGGACAGAGGTCCTGAGGATGCCCCACACACAGCAAAACATGGTGGTGGTT 2332
 DB 60 ----- 60
 QY 2333 AAACACAGGCTGGATCAGAACCCAGGCTGAGAGGGGAGCAGGTTTGGGGACCTTCC 2392
 DB 60 ----- 60
 QY 2393 TGGGGAAGGACATTTATACATGGCATGAAGGACTTGGATTTTCCAAAGGCCAAGGAAGT 2452
 DB 61 -----LeuTyr----- 62
 QY 2453 AGGCAAGGGCCCTGGAGGTGAGCTGGACTTGGCAGTGGGCGATGCAAGCCCATTTGGGCAA 2512
 DB 62 ----- 62
 QY 2513 CATATGTTATGGAGTACAAAGTCCCTTCTGTGACACACAGAGGAAGGCTTGGGAATG 2572
 DB 62 ----- 62
 QY 2573 GAAGATGAGTTAGTCTCTGAGTCCGTTTAAATCAGAAATCGAGGATGAAGGGGTGCAG 2632
 DB 62 ----- 62
 QY 2633 TGACCCGGTTCAACCTTTTGACACTGTGGGTCTCTGGGCTCACTGTCTACCGGCATGA 2692
 DB 62 ----- 62
 QY 2693 CCATCATCTGGGAATGGGATGCTAACTGGGGCTCTCGGCAATTTTGGTACTCTTGCAA 2752
 DB 62 ----- 62
 QY 2753 GGTATACCTGGGTGAGCATCAAACTGAGTTCTCTCATCACAGAAGGTGACCCCA 2812
 DB 62 ----- 62
 QY 2813 CCCCTGCCACGATCAGSAGGCTGGTCTCTCTTCCACTGTCTCACTCTCTGGTAGCC 2872
 DB 62 ----- 62
 QY 2873 CCGGGGTCTCCAAAGTTCAATAGGACTAGGACCTGTAGTCTGGGGTGTCTGGCTT 2932
 DB 62 ----- 62
 QY 2933 GACAAGAGCCCTGACCCCTCTCTGAGTTGCGGCGCGCTTGGGGACGTGTTTCAGCC 2992
 DB 63 -----LysLeuGlnAsnArgTyrGlyAspValPheSerL 74
 QY 2993 TGCAGCTGCTTGGACGGCGGTGCTGCTCAATGGGCTGCGGCGCGTGCAGGCGCA 3052
 DB 74 euGlnMetAlaTrpLysProMetValValIleAsnGlyLeuLysAlaMetLysGluMetL 94
 QY 3053 TGGTACCGCGCGGACGACACGGCGGCGCGCGCTGCGGCCATCTACAGGCTCTGG 3112
 DB 94 euLeuThrCysGlyGluAspThrAlaAspArgProProValProIlePheGluTyrLeuG 114
 QY 3113 GCTTCGGGCGGTTCCTCCAAAGCAAGCGCGGTGGGGGACAGACCGGCTTCCGTGGG 3172
 DB 114 lyValIysProGlySer-Gln----- 120
 QY 3173 CCGCGGTGGACAGTGCAGCTAGCCCAAGCAGCGCGGACAGGGCGTGGGGTCTTGACGT 3232
 DB 120 ----- 120
 QY 3233 GAACAGAGATAAGCCAGCGAGTGGGCTGAGGACAGTGGGCCAGGAAACCACTGCAC 3292
 DB 120 ----- 120
 QY 3293 GGGGAGGTGCGAGTCTGTGGGCTGGAGGGGGGGGCTACTGCCACCGCGCCAGAA 3352
 DB 120 ----- 120
 QY 3353 GCCCGGTGGCGAGGCTGATGTCGAGTGGCGGTGGGGGACCGCGCTATGCTGCG 3412

Db 120 ----- 120
QY 3413 GCGTCAGTGTGGCGGACGGCGGGATCTTCTTGTAGTGGAAAGGTGTGAGGTGGC 3472
Db 120 ----- 120
QY 3473 AGAGACAGGTGGGCCAAACCCCGCCAGCAGCGGAGCAATGTGGTGAGCAAGAG 3532
Db 120 ----- 120
QY 3533 TGGGCCCTGTGCCAGCTGGACCGGCTAGGAGCTGCGGAGACCTTGTGGAGCGCCAGG 3592
Db 120 ----- 120
QY 3593 GTTGAGTGGGTGGCGAGGGTGGGGCCAAAGGCTTCATGGCAACGCCACGTGTCGTC 3652
Db 120 ----- 120
QY 3653 CCGCCCCCAGGGGTGATCTGTGCGCTATGGCGCGGTGGCGGAGCAGGCGCTTC 3712
Db 121 ----- GlyValValLeuAlaProTyrGlyProGluTrpArgGluGlnArgArgPhe 137
QY 3713 TCGGTGTCCACCTTCGCAACTTGGCCCTGGCAAGAAAGTGCCTGGAGCAGTGGGTGACC 3772
Db 138 SerValSerThrLeuArgAsnPheGlyLeuGlyLysLysSerLeuGluAspTrpValThr 157
QY 3773 GAGGAGCGCGCTGCTTTGTGCCCTTCGCGACCAAGCCGGTGGGTGATGGGCAGAG 3832
Db 158 LysGluAlaAsnHisLeuCys ----- Asp-AlaPheTh 168
QY 3833 GGCACAAAGCGGAAGTGGGAAGGCGGGGACGAGAGAGCAACCCCTTACCCGATCTC 3892
Db 168 rAlaGlnAla ----- 171
QY 3893 CCCACCCCGAGACGCCCTTTCGCGCCCAACGGCTCTTTGACAAAGCGTGAGCAAGT 3952
Db 172 ----- GlyGlnProIleAsnProAsnProMetLeuAsnLysSerThrCysAsnVa 188
QY 3953 GATCGCTCTCTACCTCGGGCGCGCTTCGAGTACGACGACCTCTCTCTCAGGCT 4012
Db 188 IleAlaSerLeuIlePheAlaArgPheGluTyrGluAspProPheLeuIleArgMe 208
QY 4013 GCTGACCTAGCTCAGAGGAGCTGAAGGAGAGTGGGCTTTCGCGGAGGTGCGGAG 4072
Db 208 tLeuLysValLeuGluGlnSerLeuThrGluValSerGlyLeuIleProGlu ----- 225
QY 4073 CGAGACCGAGGAGTCTCTGCGGGCGAGCTCTCTGAGAGGTGCGGGGTGACTGGGG 4132
Db 225 ----- 225
QY 4133 CCTCCGAAGGCGAGGATTGTGCATAGATGGGTTTGGGAAGGACATTCAGGAGACCCAC 4192
Db 225 ----- 225
QY 4193 TGTAAAGGGCCTGGAGGAGGAGGGACATCTCAGACATGGTCTGGGAGAGGTGTC 4252
Db 225 ----- 225
QY 4253 CGGGTCAGGGGACACGAGAGGCCAAGGACTCTGTATCCCGCTCCACGTGGAGATT 4312
Db 225 ----- 225
QY 4313 CGATTTTAGTTTCTCTCTGGGCAAGAGAGAGGGTGGAGCTGGCACTTGGGAGGGA 4372
Db 225 ----- 225
QY 4373 CTTGTGAGGTGAGTGAAGACAGGAGGCGCCCTTACCTGGAGATGGCTGGGGC 4432
Db 225 ----- 225
QY 4433 CTGAGACTTGTCCAGGTGAACGACAGACAGGAGGATTGAGACCCCGTTCTGTCTGGT 4492

Db 225 ----- 225
QY 4493 GTAGGTGCTGAATGCTGTCTCCCGTCTCTCTGCAATCCACGCGTGGCTGGCAAGGTCT 4552
Db 226 ---ValLeuAsnAlaPheProIleLeuLeuArgIleProArgLeuAlaAspLysAlaLe 244
QY 4553 ACCTTCCAAAGCTTCTCTGACCCAGCTGATGCTGCTAACTGAGACACAGGATCAC 4612
Db 244 uGlnGlyLysSerPheIleAlaIleLeuAsnLeuLeuThrGluAsnArgThrTh 264
QY 4613 CTGGAGCCACCCAGCACCAGCAGCTGAGGCTTCTCTGGCAAGAGAGAGAA 4672
Db 264 rTIPAspProValGlnAlaProArgAsnLeuThrAspAlaPheLeuAlaGlnIleGly 284
QY 4673 GGTGAGAGTGGCTGCCACGCTGGGGGCAAGGGTGGTGGTGAACGTCCAGGAGGAAT 4732
Db 284 s ----- 284
QY 4733 GAGGGAGGCTGGCAAAAGGTTGGACCGATGCATCACCCGCGAGCGCATCTGGGCTG 4792
Db 284 ----- 284
QY 4793 ACAGGTGCAGAAATTGGAGGTCAATTGGGGGCTACCCCGTTCTATCCCTGAGTATCTCT 4852
Db 284 ----- 284
QY 4853 CGGCCCTGCTCAGGCCCAAGGGAGCCCTGAGAGCAGCTTCAATGATGAGAACCTCGCGAT 4912
Db 285 ---AlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuLeuMe 300
QY 4913 AGTGTGGGTAACTGTCTTCCTGCGGAGTGGTGCACCTCGACCTCGACCGCTGGCTGGG 4972
Db 300 tValValArgAspLeuPheGlyAlaGlyMeLeuThrThrSerThrLeuSerTrpAl 320
QY 4973 CCTCTGCTCATGATCTCTACCTCTGATGTGCAGCTGAGCCAGCTGGGGCCCAAGGCA 5032
Db 320 aLeuMetLeuMetIleLeuHisProAspValGln ----- 331
QY 5033 GGGACTGAGGAGGAAGGGTACAGCTGGGGGCCCTGGGCTTAGTGGGACACCGGGGC 5092
Db 331 ----- 331
QY 5093 TTCCAGCACAGCGCTGGCCAGCTCTCTAAGCCTTAATCTCTCCACACAGGAGGAGG 5152
Db 331 ----- 331
QY 5153 AGAGTGTCCCTGGGTGTGTGACCCATTGTGGGGAGCATGTCTGTCCAGTCTCGTCCAA 5212
Db 332 ---Arg-ArgValGln 335
QY 5213 CAGGAGATCGACGACGTGATAGGGCAGGTGGGGCACACAGAGATGGGTGACAGGCTCAC 5272
Db 336 GlnGluIleAspGluValIleGlyValArgHisProGluMetAlaAspGlnAlaHis 355
QY 5273 ATGCCCTACACCTGCGCTGATTACAGAGTGCAGCGCTTTGGGGACATCATCCCGCTG 5332
Db 356 MetProTyrThrAsnAlaValIleHisGluValGlnArgPheGlyAspIleValProVal 375
QY 5333 AGTGTGACCCATATGACATCCGTGACATCGAAGTACAGGCTTCGCGCATCCCTAAGGTA 5392
Db 376 AsnLeuProArgIleThrSerHisAspIleGluValGlnAspPheLeuIleProLys --- 394
QY 5393 GGCCTGGCGCCCTCTCTACCCAGCTCAGCACCGACCTGGTGTGATGCCCCAGCATGGC 5452
Db 394 ----- 394
QY 5453 TACTGCCAGGTGGGCCACTCTAGGAACCTGTGGCCACTAGTCTCAATGCCACCACT 5512
Db 394 ----- 394
QY 5513 GACTGTCCCACTTGGGTGGGGGTCCAGAGTATAGGAGGGCTGGCGCTGTCCATCCAGA 5572
Db 394 ----- 394

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: February 25, 2004, 02:16:18 ; Search time 147.885 Seconds
(without alignments)
13370.622 Million cell updates/sec

Title: US-09-820-788A-3
Perfect score: 18731
Sequence: 1 agccttacaagtctggga.....ccagggtcagtcggcaggt 10278

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DRV=xlp
-Q/cgn2 1/USPTO_spool_p/US09820788/runat_24022004_141427_20051/app_query.fasta_1.12174
-DB=PIR78 -QPMF=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=us09820788 @cgn 1.1 292 @runat_24022004_141427_20051 -NCPUP=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	1795	9.6	497 1 O4HUD1
2	1675	8.9	497 1 G02938
3	1318.5	7.0	500 1 S37284
4	1311.5	7.0	500 1 JE0258
5	1280.5	6.8	500 1 D31579
6	1268.5	6.8	500 1 JCS819
7	1252.5	6.7	500 1 JE0259
8	1249.5	6.7	500 1 JCA157
9	1176	6.3	500 1 S16872
10	1175.5	6.3	500 1 B26822
11	1161.5	6.2	500 1 JC4153
12	1146	6.1	504 1 O4RTD5
13	1137	6.1	504 1 A27384
14	1131	6.0	504 1 I49428

15	1126	6.0	504	1	A26822	debrisoquine 4-hyd
16	1122	6.0	504	2	I49427	cytochrome P450 16
17	1119	6.0	504	1	A30247	cytochrome P450 2D
18	1034.5	5.5	505	1	S19169	cytochrome P450 2D
C 19	1690	3.0	1690	1	CGHU1B	collagen alpha 4(I
C 20	520	2.8	3570	2	T45025	mucin MUC5B, trach
21	507.5	2.7	294	1	B27384	probable truncated
C 22	499.5	2.7	2715	2	T13049	eyelid - fruit fly
23	495	2.6	1690	1	CGHU1B	collagen alpha 4(I
C 24	489	2.6	4957	2	T03455	ALR protein - huma
25	486.5	2.6	2944	2	A54849	collagen alpha 1(V
26	484.5	2.6	13288	2	T03099	mucin, submaxillar
27	481	2.6	1466	1	CGHU7L	collagen alpha 1(I
C 28	477.5	2.6	2944	1	A54849	collagen alpha 1(V
C 29	471	2.5	1691	1	S22917	collagen alpha 5(I
C 30	467.5	2.5	213	2	T47135	hypothetical prote
C 31	459.5	2.5	2142	2	B35098	MHC class III hist
32	459	2.5	1691	1	S22917	collagen alpha 5(I
33	457	2.4	501	1	A04938	cytochrome P450 1b
C 34	456.5	2.4	5262	2	T03454	ALR protein - huma
C 35	455.5	2.4	1453	2	S21626	collagen alpha 2(I
36	455.5	2.4	1763	2	S16366	collagen alpha 2(I
C 37	449	2.4	1763	2	S16366	collagen alpha 2(I
38	448.5	2.4	1464	2	S59856	collagen alpha 1(I
C 39	448.5	2.4	1670	1	CGHU3B	collagen alpha 3(I
C 40	446.5	2.4	1669	1	CGMS4B	collagen alpha 1(I
C 41	446	2.4	1188	2	S49915	extensin-like prot
42	445.5	2.4	1049	1	CGBO7S	collagen alpha 1(I
43	445.5	2.4	1549	2	I48103	type VII collagen
44	443	2.4	1806	1	CGHU1E	collagen alpha 1(X
C 45	442.5	2.4	1464	2	S59856	collagen alpha 1(I

ALIGNMENTS

RESULT 1

OAHU01
debrisoquine 4-hydroxylase (EC 1.14.14.-) cytochrome P450 2D6 - human
N/Alternate names: CYP2D6; cytochrome P450 isozyme 2D; cytochrome P450db1
C/Species: Homo sapiens (man)
C/Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text_change 03-Mar-2000
C/Accession: S01199; A28883; JC4156; A33629; A30335
R/Gonzalez, F.J.; Skoda, R.C.; Kimura, S.; Umeno, M.; Zanger, U.M.; Nebert, D.W.; Gelbo
Nature 331, 442-446, 1988
A/Title: Characterization of the common genetic defect in humans deficient in debrisoqu
A/Reference number: S01199; PMID:88122614; PMID:3123997
A/Accession: S01199
A/Status: translation not shown
A/Molecule type: mRNA
A/Residues: 1-497 <CON>
A/Cross-references: EMBL:X08006; NID:G30450; PIDN:CAA30807.1; PID:G30451
R/Gonzalez, F.J.; Vilbois, F.; Hardwick, J.P.; McBride, O.W.; Nebert, D.W.; Gelboin, H.
Genomics 2, 174-179, 1988
A/Title: Human debrisoquine 4-hydroxylase (P450ID1): cDNA and deduced amino acid sequ
A/Reference number: A28883; PMID:88314109; PMID:3410476
A/Accession: A28883
A/Molecule type: mRNA
A/Residues: 1-497 <CON>
A/Cross-references: EMBL:M20403; NID:G181349; PIDN:AAA52153.1; PID:G181350
R/Jiang, Q.; Voigt, J.M.; Colby, H.D.
Biochem. Biophys. Res. Commun. 209, 1149-1156, 1995
A/Title: Molecular cloning and sequencing of a guinea pig cytochrome P4502D (CYP2D16):
A/Reference number: JC4153; PMID:95251703; PMID:7733969
A/Accession: JC4156
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-497 <JIA>
R/Kimura, S.; Umeno, M.; Skoda, R.C.; Meyer, U.A.; Gonzalez, F.J.
Am. J. Hum. Genet. 45, 889-904, 1989
A/Title: The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification
A/Reference number: A33629; PMID:90072069; PMID:2574001
A/Accession: A33629
A/Molecule type: DNA

A:Residues: 1-373,'V',375-497 <KIM>
A:Cross-references: EMBL:M33388; NID:g181303; PIDN:AAA53500.1; PID:g181304
R:Manns, M.P.; Johnson, E.F.; Griffin, K.J.; Tan, E.M.; Sullivan, K.F.
J. Clin. Invest. 83, 1066-1072, 1989
A:Title: Major antigen of liver kidney microsomal autoantibodies in idiopathic autoimmune
A:Reference number: A30335; MUID:89155788; PMID:2466049
A:Accession: A30335

A:Molecule type: mRNA
A:Residues: 125-373,'V',375-485,'T',487-497 <MAN>
A:Cross-references: EMBL:M24499; NID:G522194; PIDN:AAA36403.1; PID:G522195
C:Genetics:

A:Gene: GDB:CYP2D6
A:Cross-references: GDB:132127; OMIM:124030
A:Map position: 22q13.1-22q13.1
A:Introns: 60/3; 118/1; 169/1; 222/3; 281/3; 329/1; 391/3; 439/1
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: Chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F:302-465/Domain: cytochrome P450 homology <CYP>
F:443/Binding site: heme iron (Cys) (axial ligand) #status predicted

Alignment Scores:

Pred. No.: 2.43e-83 Length: 497
Score: 1795.00 Matches: 464
Percent Similarity: 34.14% Conservative: 14
Best Local Similarity: 33.14% Mismatches: 19
Query Match: 9.58% Indels: 903
DB: 1 Gaps: 9

US-09-820-788a-3 (1-10278) x O4HUD1 (1-497)

QY	2078	ATGGGGCTAGAGACACTGGTCCCTGGCCATGATAGTGGCCATCTTCTGCTCTGCTG	2137
DB	1	MetGlyLeuGluAlaLeuValProLeuAlaValIleAlaIlePheLeuLeuVal	20
QY	2138	GACCTGATGACCGGCACCAAGCTGGGTGACGCTACCCGCCAGGTCCCTGCACTG	2197
DB	21	AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProGlyProLeu	40
QY	2198	CCGGGGCTGGGCAACTGCTGCTGATGAGACTTCAGAACACACATCTGCTGACCA	2257
DB	41	ProGlyLeuGlyAsn-LeuLeuHisValAspPheGlnAsnThrProTyrCysPheAspG1	60
QY	2258	GGTGAGGAGGAGGTCTGGAGGGCGGCAGAGGTCTGAGGATGCCACCACAGCAA	2317
DB	60	n-----	60
QY	2318	CATGGGTGGTGGTTAAACACAGGCTGGATCAGAAGCCAGGCTGAGAAGGGGAAGCAGG	2377
DB	60	-----	60
QY	2378	TTTGGGGGACGTTCCTGGGGAAGGACATTTATACATGGCATGAGACTGGATTTTCAA	2437
DB	60	-----	60
QY	2438	AGCCCAAGAGAGTAGGCAAGGCCCTGGAGGTGGAGCTGGACTTGGCAGTGGGCATGC	2497
DB	60	-----	60
QY	2498	AAGCCCAATTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACACCAAGGA	2557
DB	60	-----	60
QY	2558	AAGGCTTTGGGAATGGAAGATGAGTTAGTCTGAGTGGCGTTTAAATCAGGAATCGAGG	2617
DB	60	-----	60
QY	2618	ATGAAGGGGTGACGTGACCGGTTCAAACCTTTTTCACCTGTTGGGTCTCGGGCCTCACT	2677
DB	60	-----	60
QY	2678	GCTACCGGCATGGACCATCATCTGGGAATGGGATGCTAACTGGGGGCTCTCGGCAATTT	2737
DB	60	-----	60

QY	2738	TGGTGACTCTTGCAGAGTCAATACCTGGGTGAGCATCCAAACTGAGTTCCTCCATCAG	2797
DB	60	-----	60
QY	2798	AAGGTGTGACCCCAACCCCTGCCACGATCAGGAGGTGGGTCTCTCTCCCTTCCACCTGC	2857
DB	60	-----	60
QY	2858	TCACTCTCTGTAGCCCGGGGTGTCOAAGTTCAATAGGACTAGGACCTGTAGTCTG	2917
DB	60	-----	60
QY	2918	GGTGATCTCTGGCTTGACAAAGGCCCTGACCCCTCTCTGAGTTGGGGCGCGCTTCG	2977
DB	61	-----Leu-ArgArgArgPheG	66
QY	2978	GGGACGTCTTCAGCTGCTGAGCTGGCGCGGTGGTCTCAATGGGCTGGCGG	3037
DB	66	lyAspValPheSerLeuGlnLeuAlaTrpThrProValValValLeuAsnGlyLeuAla	86
QY	3038	CCGTGCGGCGAGTGTGACCCCGCGGAGGACACGCGCGGCGCTGCGGCCA	3097
DB	86	laValArgGluAlaLeuValThrHisGlyGluAspThrAlaAspArgProValProI	106
QY	3098	TCTACAGGTCTCGGCTTCGGCGCGGTTCCTCCAAAGCAAGCGGCTGGGGGACAGAG	3157
DB	106	leThrGlnIleLeuGlyPheGlyProArgSer-Gln	117
QY	3158	CCGCGTTTTCGTGGGCGCGGTGGACAGTACCGTAGCCCAAGCGCGGACAGGGCG	3217
DB	117	-----	117
QY	3218	TGGGTCTTGACGTGAAACAGAGATAAGCCACGAGTGGGTGAGGACAGTGGGCCA	3277
DB	117	-----	117
QY	3278	GAAACACCTGACAGGGGAGGTGCGAGTCTGTGGGTGGGAGGGGGGCTACTGC	3337
DB	117	-----	117
QY	3338	CCAGACCGCCAGAAAGCCCGGTGGCGGAGGTGATCGTCGAAGTGGCGGTGGCGGGAC	3397
DB	117	-----	117
QY	3398	CGCGCTATGCTGCGGGCTCAGTGTGGCGGACGCGCGGATCTTCTTGAAGTGAAG	3457
DB	117	-----	117
QY	3458	GTGTGAGGTGGGACAGACAGGTGGGGCCAAACCCCGCCGAGGAGGAGCAATG	3517
DB	117	-----	117
QY	3518	TGGGTGAGCAAGAGTGGGCGCTGTGCCAGCTGACCGGGCTAGGACTGGGAGACC	3577
DB	117	-----	117
QY	3578	TTGTGAGGCCAGGGTTGGAGTGGGTGGCGGAGGTGGGGCCCAAGGCTTTCATGGCAAC	3637
DB	117	-----	117
QY	3638	CCCCACGTGTCCTCCCGCCCCAGGGGTGATCTCTGCGCTATGGGCCCGCTGGCGC	3697
DB	118	-----GlyValPheLeuAlaArgTyrGlyProAlaTrpArg	129
QY	3698	GAGCAGAGCGCTTCTCGCTGCCACTTGGCAACTTGGGCTGGGCAAGTGGCTG	3757
DB	130	GluGlnArgArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyIleSerLeu	149
QY	3758	GAGCAGTGGGTGACCGGAGGCGCTCTTTTGTCCGCTTCGCCGACCAAGCCGCTG	3817
DB	150	GluGlnTrpValThrGluAlaAlaCysLeuCysAlaAlaPheAlaAsn	166

```
QY 3818 GGTGATGGCAGAGGGCAAAAGCGGGAACTGGGAAGCGGGGACGGAGAAGGCAACC 3877
Db 166 ----- 166
QY 3878 CCTTACCCGATCTCCCAACCCAGAGCGCCCTTTCCGCCCAAGCGCCTCTTGACAA 3937
Db 167 -----His-Ser-----GlyArgProPheArgProAsnGlyLeuLeuAspLys 180
QY 3938 AGCCGTGAGCAACGTGATCGCTTCCCTCACCTGCGGGCGCCGCTTCGAGTACGACGACCC 3997
Db 180 sAlaValSerAsnValIleAlaSerLeuThrCysGlyA-rgPheGluTy-AspAspPr 200
QY 3998 TGCTTCTCAGCTCTCGACTAGCTCAGAGGAGCTGAAGAGAGAGTCGGGCTTTCT 4057
Db 200 oArgPheLeuArgLeuLeuAspLeuAlaGlnGluGlyLeuLysGluLysSerGlyPheLe 220
QY 4058 GCGCAGAGTGGCAGCGAGAGACCGAGGAGTCTCTCAGGGCGGAGTCTCTGAGAGTGCC 4117
Db 220 uArgGlu----- 222
QY 4118 GGGGCTGGACTGGGGCTCCGAAGGCGAGGATTTGTATAGATGGGTTTGGGAAAGGACAT 4177
Db 222 ----- 222
QY 4178 TCCAGGAGACCCACCTGTAAAGAGGCCCTGGAGAGGAGGGGACATCTCAGACATGGTCG 4237
Db 222 ----- 222
QY 4238 TGGGAGAGGTGTGCCCGGGTCAGGGGGCACAGAGAGGCCAAGGACTCTGTACCCCGT 4297
Db 222 ----- 222
QY 4298 CCACGTTGGAGATTCGATTTTAGGTTTCTCTCTGGCAAGGAGAGGGTGGAGGCTG 4357
Db 222 ----- 222
QY 4358 GCACTTGGGAGGGACTTGGTGAGGTCAGTGGTAAGGACAGGCGGCCCTCGGCTTACCT 4417
Db 222 ----- 222
QY 4418 GGAGATGGCTGGGGCTGAGACTTGTTCAGGTGAACGACAGAGACAGAGGAGATTGAGAC 4477
Db 222 ----- 222
QY 4478 CCGCTTCTGTCTGGTGTAGTGTGAATGTGTCCCGTCTCTCGCACATCCAGCGCT 4537
Db 223 -----ValLeuAsnAlaValProValLeuLeuHisIleProAlaLe 236
QY 4538 GGCTGGCAAGGCTCTACGCTTCCAAAGGCTTTCTGTGACCCAGCTGGATGAGCTGCTAAC 4597
Db 236 uAlaGlyLysValLeuArgPheGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGAGCACAGGATGACCTGGGACCCAGCCAGCCAGCCAGACCTCGACTGAGCGCTTCTCT 4657
Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GGCAAAGAGGAGAGAGTGCAGATGCTGCCAGGTGGGGGCAAGGGTGGTGGTGA 4717
Db 276 uAlaGluMetGluLys----- 281
QY 4718 CGTCCAGGAGGAATGAGGGGAGGCTGGGCAAAAGGTTGGACCAAGTGATCATCCCCGCGA 4777
Db 281 ----- 281
QY 4778 GCGGCATCTGGGCTGACAGGTGCAGAAATTGGAGGTCAATTGGGGGCTACCCGTTCTATC 4837
Db 281 ----- 281
QY 4838 CCTGTAGTATCTCTCGGCCCTGCTCAGGCCAAGGGGAGCCCTTGAGAGCAGCTTCAATGA 4897
Db 282 -----AlaLysGlyAsnProGluSerPheAsnAs 292
QY 4898 TGAGAACCTCGCATAGTGTGGGTAACTGTTCCTTTCGCGGATGGTGACCACTCGAC 4957
```

```
Db 292 pGluAsnLeuArgIleValValAlaAspLeuPheSerAlaGlyMetValThrThrSerTh 312
QY 4958 CACGCTGGGCTGGGCGCTCTCTCATGATCTACACCTGGATGTGCAGCGTGAGCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuLeuMetIleLeuHisProAspValGln----- 328
QY 5018 CTGGGGCCCAAGCAGGAGGACTGAGGGAGGAAGGTACAGCTGGGGGCCCTCTGGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCCGGGGCTTCCAGACAGGCGTGCCAGGCTCTGTAAAGCCTAATCTCTCCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGGAGGAAGGAGAGTGTCCCTGGGTCTGACCCATTTGTGGGGACGCATGTCTGT 5197
Db 328 ----- 328
QY 5198 CCAGTCCGTGTCCAAACAGGAGATCGACGACGTGATAGGCGAGGTGGCGACAGAGATG 5257
Db 329 -Arg--Arg-ValGlnGlnGluIleAspAspValIleGlyGlnValArgArgProGluMet 347
QY 5258 GGTGACGAGGCTCACATGCCCTACACCACTGCCGTGATTCACAGAGTGCAGCGCTTTGG 5317
Db 348 GlyAspGlnAlaHisMetProTy-rrThrAlaValIleHisGluValGlnArgPheGly 367
QY 5318 GACATCATCCCTCCTGAGTGTGACCATATGACATCCGTGACATCGAAGTACAGGCTTTC 5377
Db 368 AspIleValProLeuGlyMetThrHisMetThrSerArgAspIleGluValGlnGlyPhe 387
QY 5378 CGCATCCCTTAAGTAGGCGCTGCGGCCCTCTCACCCAGCTCAGCACACGACCTGGTGA 5437
Db 388 ArgIleProLys----- 391
QY 5438 TAGCCCCAGCATGGCTACTGTGCCAGGTGGGCCCACTCTAGGAACCCCTGGCCACCTAGTCTCT 5497
Db 391 ----- 391
QY 5498 CAATGCCACCACACTGACTGTCTCCCACTTGGGTGGGGGTCCAGAGATAGGCAGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTCCATCCAGAGCCCCCGTCTAGTGGGGGAGACAAACAGGACCTGCCAGATGTTG 5617
Db 391 ----- 391
QY 5618 GAGGACCCAGCGCTGACAGGAGAGGGGGCAGTGTGGGTGCTCTGTAGAGAGGTGACTGC 5677
Db 391 ----- 391
QY 5678 GCGCTGTGTGGGTCCGAGAGGGGTACTGTGGAGCTTCTCGGGGCGCAGGACTAGTTGACA 5737
Db 391 ----- 391
QY 5738 GAGTCCAGCTGTGTGCCAGCAGTGTGTCTCCCGTGTGTTTGGTGGCAGGGGTCCCAG 5797
Db 391 ----- 391
QY 5798 CATCTAGAGTCACTGCCCACTCTCAACCTGCATCTCTGTCGCCAGGGAACGACACTCAT 5857
Db 392 -----Gly--ThrThrLeuIle 396
QY 5858 CACCAACCTGTCTCATCGGTGCTGAAGATGAGCGCTCTGGGAGAGCCCTTCGCTCCCA 5917
Db 396 eThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHi 416
QY 5918 CCCCAGAACACTTCTCGGATGCCAGGCGCACCTTGTGAAGCCGAGGCGCTTCTCGCTTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPh 436
QY 5978 CTCAGCAGGTGCTGTGGGGAGCCCGGCTCTCTGTCCTCCCTTCCGTGGAGTCTTTGAGGGG 6037
```

Db 436 e--Ser----- 437
QY 6038 TATACCCAGGAGCAGGCTACTGACGCCCTCCCTCCCAAGGCGCGTGCATGC 6097
Db 438 -----AlaGlyArgGalaCys 443
QY 6098 CTCGGGAGCCCTGGCCCGCATGGAGCTTCTCTTCTTCCACCTCCCTGCTGCAGCAC 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPhePheThrSerLeuLeuGlnHis 463
QY 6158 TTCAGTCTTCGTCGCGCCCGGACAGCCCGCCAGCACTCTGTCGTCAGCTTT 6217
Db 464 PheSerPheSerValProThrGlyGlnProArgProSerHisHisGlyValPheAlaPhe 483
QY 6218 CTGGTGACCCATCCCTTACAGCTTTGTGTGTGTGTCGCCCGC 6259
Db 484 LeuValSerProSerProTyrGluLeuCySAIaValProArg 497

RESULT 2
G02938
probable debrisoquine 4-hydroxylase (EC 1.14.14.-) cytochrome P450 - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Mar-2000
C:Accession: G02938
R:Lawton, M.P.; Laddison, K.J.; Speirs, A.A.; Mankowski, D.C.; Tweedie, D.J.
submitted to the EMBL Data Library, October 1995
A:Reference number: G12616
A:Accession: G02938
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-497 <LAW>
A:Cross-references: EMBL:U38218; NID:g1022899; PIDN:AAA79722.1; PID:g1022900
C:Genetics:
A:Gene: CYP2D17
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: Chromoprotein; heme; iron; metalloprotein; monooxygenase; oxidoreductase; tyrosine hydroxylase
F:302-465/Domain: cytochrome P450 homology <P45>
F:443/Binding site: heme iron (Cys) (axial ligand) #status predicted

Alignment Scores:
Pred. No.: 2,82e-77 Length: 497
Score: 1675.00 Matches: 442
Percent Similarity: 33.00% Conservative: 20
Best Local Similarity: 31.57% Mismatches: 35
Query Match: 8.94% Indels: 903
DB: 1 Gaps: 8

US-09-820-788A-3 (1-10278) x G02938 (1-497)

QY 2078 ATGGGGCTAGAGCACTGGTGGCCCTGGCCATGATAGTGGCCATCTTCTGCTCTGCTG 2137
Db 1 MetGluLeuAspAlaLeuValProLeuAlaValThrValAlaIlePheLeuLeuVal 20
QY 2138 GACCTGATGACCGGCACCAACGCTGGGCTGACGCTACCCGCCAGGTCCCTGCACTG 2197
Db 21 AspLeuMetHisArgArgGlnArgTrpAlaAlaArgTyrProProGlyProLeuProLeu 40
QY 2198 CCCGGCTGGGCAACCTTCTGCTCATGTGGACTTCAGAACACACCATACTGCTCGACCA 2257
Db 41 ProGlyLeuGlyAsn-LeuLeuHisValAspPheLysAsnThrProTyrCysPheAspG1 60
QY 2258 GTGAGGAGGAGGTCTCTGGAGGGCGGACAGGTCTCTGAGGATGCCCCACCACAGCAA 2317
Db 60 n----- 60
QY 2318 CATGGTGTGGGTTAAACCAACAGCTGGATCAGAAGCCAGGCTGAGAAGGGGAGCAGG 2377
Db 60 ----- 60
QY 2378 TTTGGGGACGTTCTCTGGGGAAGGACATTTATACATGGCATGAAGGACTGGATTTTCAA 2437
Db 60 ----- 60

QY 2438 AGGCCAAGGAAGAGTAGGGCAAGGCCCTGGAGGTGGAGCTGGACTTGGCAGTGGGCATGC 2497
Db 60 ----- 60
QY 2498 AAGCCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACACAGAAGGA 2557
Db 60 ----- 60
QY 2558 AAGCCCTTGGGAATGGAAGATGAGTTAGTCTGAGTGGCGTTTAAATCAGAAATCGAGG 2617
Db 60 ----- 60
QY 2618 ATGAAGGGGGTGCAGTGACCCGGTTCAAACCTTTTGGCACTGTGGGTCTCTCGGGCCTCACT 2677
Db 60 ----- 60
QY 2678 GCTCACCGGCATGGACCATCATCTGGGAATGGGATGCTAACTGGGGGCTCTCGGCAATTT 2737
Db 60 ----- 60
QY 2738 TGGTGACTCTTGGCAAGGTACATACCTGGGTGACGATCCAAACTGAGTTCTCTCCATCAG 2797
Db 60 ----- 60
QY 2798 AAGGTGTGACCCCAACCCCTGCCACGATCAGGAGGCTGGGTCTCTCTCTTCCACCTGC 2857
Db 60 ----- 60
QY 2858 TCACCTCTGTAGCCCGGGGTCTCCAAAGTTTCAAATAGGACTAGGACCTGTAGTCTG 2917
Db 60 ----- 60
QY 2918 GGGTGATCTCGCTTGACAAAGGCCCTGACCTCCCTCTCGAGTTGGCGGCCCTTCG 2977
Db 61 -----Leu-ArgArgArgPhe 66
QY 2978 GGGACGCTTTCAGCTGACCTGCGCTGGCGCGGTGGTCTGCTGCTCAATGGGCTGGCGG 3037
Db 66 LysnValPheSerLeuGlnLeuAlaIleThrProValValLeuLeuGlyLeuAla 86
QY 3038 CCCTGCGGAGCGATGCTGACCCCGCGGAGGACACGCGCGACGCGCCGCTTGGCGCCA 3097
Db 86 LaValArgGluAlaLeuValThrCysGlyGluAspThrAlaAspArgProProValProI 106
QY 3098 TCTACAGGTCTCTGGGCTTCCGGCGGTCTCCCAAGCAAGCGCGGTGGGGGACAGAGA 3157
Db 106 LeAsnGlnValLeuGlyPheGlyProArgSer-Gln----- 117
QY 3158 CCGGTTTTTCGTGGGCCCCGGGTGGACAGTACCGTAGCCCAAGCAGCGCCGACAGGCG 3217
Db 117 ----- 117
QY 3218 TGGGCTCTGGACGCTGAACAGAGATAAGGCCAGCGAGTGGGTGGTGGGACAGTGGGCGCA 3277
Db 117 ----- 117
QY 3278 GAAACACACTGCACGGGGAGGTGCGAGTCTGTGGGTGGAGGGGGCGGGGCTACTGCG 3337
Db 117 ----- 117
QY 3338 CCAGACCGCCAGAACCCCGTGGCGAGGCTGATGCTCGAAGTGGCGGTGGCGGGGAC 3397
Db 117 ----- 117
QY 3398 CGCGCTATGCTGGGGCTCAGTGTGGCGGGAACCGGCGGGATCTTCTTGAAGTGAAG 3457
Db 117 ----- 117
QY 3458 GTGGTACAGGTGGGACAGACAGAGTGGGGGCCAAACCCCGCCCGGAGGAGCAATG 3517
Db 117 ----- 117
QY 3518 TGGGTGACAAAGAGTGGGCCCTGTGCCAGCTGACCGGCTAGGAGCTGGCGGAGACC 3577

Db 117 ----- 117
QY 3578 TTGTGAGGCCAGGGTTGGAGTGGGTGGCGAGGGTGGGGCCAAAGGCTTCATGGCAAC 3637
Db 117 ----- 117
QY 3638 GCCACGTGTCCGTCCGCCCCACAGGGGTATCTCTCGCGCTATGGGCCCGCGTGGCGC 3697
Db 118 -----
GlyValPheLeuAlaArgTyrGlyProAlaTrpArg 129
QY 3698 GACGACAGGGCTTCCCGTGTCCACTTGGCGCACTTGGGCTGGGCAAGAGTGGCTG 3757
Db 130 GluGlnArgArgPheSerValSerThrLeuArgAsnLeuGlyLeuGlyLysSerLeu 149
QY 3758 GACGAGTGGGTGACGAGGCGCGCTGCTTGTGTGGCGCTTCGCGCAAGCGCGTG 3817
Db 150 GluGlnTrpValThrGluAlaAlaCysLeuCysAlaAla ----- 163
QY 3818 GGTGATGGGCAGAGGGCACAAGCGGGAACCTGGGAAGCGGGGACGGAGAAGGCAACC 3877
Db 163 ----- 163
QY 3878 CCTTACCCGATCTCCGACCCCCACGAGCGCCCTTTCGCCCAACGGGCTCTTGGACAA 3937
Db 164 ----- Phe-ThrAspGlnAlaGlyArgProPheArgProAsnSerLeuLeuAsp 180
QY 3938 AGCCGTGAGCAACGTGATCGCTCCCTCACCTCGCGGCGCCGCTTCGAGTACGACGACC 3997
Db 180 sAlaValSerAsnValIleAlaSerLeuThrTyrGlyArgPheGluTyrAspAspPr 200
QY 3998 TCGCTTCTCAGGCTCTCGACCTAGCTCAGGAGGGACTGAAGGAGAGTTCGGGCTTTCT 4057
Db 200 oArgPheLeuArgLeuPheAspLeuThrHisGluAlaLeuLysGluGluSerGlyPheLe 220
QY 4058 GCGGAGGTGCGGAGGAGGACCGAGGAGTCTCTCGAGGGCGAGCTCTGAGAGTGCC 4117
Db 220 uArgGlu ----- 222
QY 4118 GGGGCTGGAGTGGGGCTCCGAAGGCGAGATTGTCATAGATGGGTTTGGAAAGGACAT 4177
Db 222 ----- 222
QY 4178 TCCAGGAGACCCCACTGTAAAGAGGCGCTGGAGGAGGGGACATCTCAGACATGGTCG 4237
Db 222 ----- 222
QY 4238 TGGGAGAGGTGTCCCGGGTCAGGGGGCACAGGAGAGGCCAAGGACTCTGTACCCCGT 4297
Db 222 ----- 222
QY 4298 CCAGGTTGGAGATTTCGATTTTAGGTTTCTCTCTGGGCAAGAGAGGGTGGAGGCTG 4357
Db 222 ----- 222
QY 4358 GCATTTGGGAGGGACTTGTGAGGTCAGTGGTAAGGACAGGCGGCCCTGGGTCTACCT 4417
Db 222 ----- 222
QY 4418 GGAGATGGCTGGGGCTGAGACTTGTTCAGGTGAACGCAGAGCACAGGAGGATTGAGAC 4477
Db 222 ----- 222
QY 4478 CCGGTTCTGTGTGTAGTGTGAATGCTGTCCCGTCTCTCTGTCACATCCAGCGCT 4537
Db 223 -----
ValLeuAsnAlaIleProLeuLeuArgIleProGlyLe 236
QY 4538 GGTGTCAGAGTCTTACGCTTCCAAAAGGCTTTCCTGACCCAGCTGAGTGTGCTAAC 4597
Db 236 uAlaGlyLysValLeuArgSerGlnLysAlaPheLeuThrGlnLeuAspGluLeuLeuTh 256
QY 4598 TGAGCACAGGATGACCTGGGACCCAGCCAGCCACCCGAGACCTGACTGAGCGCTTCTCT 4657
|||||

Db 256 rGluHisArgMetThrTrpAspProAlaGlnProProArgAspLeuThrGluAlaPheLe 276
QY 4658 GGCAAGAAGAGAGAGGTGAGAGTGGCTGCCACGCTGGGGGGCAAGGGTGGTGGGTTGAA 4717
Db 276 uAlaGluMetGluLys ----- 281
QY 4718 CGTCCAGGAGGAATGAGGGGAGGCTGGGCAAAAGTTGGACCAGTGCATCACC CGCGCA 4777
Db 281 ----- 281
QY 4778 GCCGCATCTGGGCTGACAGGTGCAGAAATTGGAGGTCAATTTGGGGGTACCCCGTTCTATC 4837
Db 281 ----- 281
QY 4838 CCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCAGCTTCAATGA 4897
Db 282 -----
AlaLysGlyAsnProGluSerSerPheAsnGI 292
QY 4898 TGAGAACCTGCGCATAGTGGGTAACTGTCTCTTCTCGGGGATGGTGACCACTCGAC 4957
Db 292 uGluAsnLeuArgMetValValAlaAspLeuPheSerAlaGlyMetValThrSerTh 312
QY 4958 CACGCTGGCGCTCGGCTCTCTCATGATCTACCTGATGTCAGCGTGACGCTGAGCCAG 5017
Db 312 rThrLeuAlaTrpGlyLeuLeuMetIleLeuHisProAspValGln ----- 328
QY 5018 CTGGGGCCCAAGGCGGACTGAGGAGGAAGGTACAGCTGGGGGCCCCCTGGGCTTAGC 5077
Db 328 ----- 328
QY 5078 TGGGACACCCGGGGTTCAGCACAGGCGTGGCCAGGCTCTGTAAAGCCTTAATCTCTCC 5137
Db 328 ----- 328
QY 5138 AACACAGGAGGAGGAGAGTGTCCTCGGTGCTGACCCATTGTGGGGAGCGATGCTGT 5197
Db 328 ----- 328
QY 5198 CCAGTCCGTGTCCAAACAGGAGATCGACGAGTGTAGGAGCGTGGCGCACAGAGATG 5257
Db 329 -Arg-ArgValGlnGlnGluIleAspValIleGlyGlnValArgProGluMet 347
QY 5258 GGTGACCAAGGCTCAATGCCCTACACCTGCGGTATTTCAGAGTGCAGCGCTTTGGG 5317
Db 348 GlyAspGlnAlaArgMetProTyrThrAlaValIleHisGluValGlnArgPheGly 367
QY 5318 GACATCATCCCTGAGTGTGACCCCATATGACATCCCGTGACATCGAAGTACAGGCTTC 5377
Db 368 AspIleValProLeuGlyValThrHisMetThrSerArgAspIleGluLeuGlnGlyPhe 387
QY 5378 CGCATCCCTAAGGTAGGCTGGCGCTCTCTCACCCAGCTCAGCACACGACCTGGTGA 5437
Db 388 LeuIleProLys ----- 391
QY 5438 TAGCCCCCAGCATGGTACTGCCAGGTGGGCCCACTCTAGGAACCTTGGCCACCTAGTCTC 5497
Db 391 ----- 391
QY 5498 CAATGCCACCACACTGACTGTCCCCACTTGGGTGGGGGTCCAGAGTATAGGAGGGCTG 5557
Db 391 ----- 391
QY 5558 GCCTGTTCATCCAGAGCCCCCGTCTAGTGGGGGAGACAAACAGGACCTGCCAGATGTTG 5617
Db 391 ----- 391
QY 5618 GAGGACCCAGCGCTCGAGGAGAGGGGCGAGTGTGGTGCCTCTGAGAGGTGTGATGC 5677
Db 391 ----- 391
QY 5678 GCCCTGCTGTGGGGTTCGGAGAGGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACA 5737
Db 391 ----- 391

```

QY 5738 GAGTCAGCTGTGTCAGGAGTGTGTGTCCTCCCGTGTGTTGCTGGCAGGGTCCAG 5797
Db 391 -----
QY 5798 CATCTAGAGTCCAGTCCCACTCTCACCTGCATCTCTGCTCCAGGAACGACACTCAT 5857
Db 392 -----Gly-ThrThrLeuPh 396
QY 5858 CACCAACTGTGTCATCGGTCTGAAGATGAGCGCTCTGGGAGAGCCCTCCGCTTCCA 5917
Db 396 eThrAsnLeuSerSerValLeuLysAspGluAlaValTrpGluLysProPheArgPheHi 416
QY 5918 CCCCAGCACTTCTCGGTGTCAGGCGCACTTCTGAGCGGAGGCTCTCTGCTT 5977
Db 416 sProGluHisPheLeuAspAlaGlnGlyHisPheValLysProGluAlaPheLeuProPh 436
QY 5978 CTCAGCAGGTGCTGTGGGAGCGCGGTCTCCGTGCTCCCTTCCGTGGAGTCTTGCAGGG 6037
Db 436 e--Ser----- 437
QY 6038 TATACCCAGGAGCCAGGCTACTGACGCCCTCCCTCCCAAGCGCGCGTGCATGC 6097
Db 438 -----AlaGlyArgArgAlaCys 443
QY 6098 CTCGGGGAGCCCTGGCGGATGAGCTTCTCTTCTTTCACCTCTCTGCTGCGAGCAC 6157
Db 444 LeuGlyGluProLeuAlaArgMetGluLeuPhePheThrCysLeuLeuGlnArg 463
QY 6158 TTCAGCTTCTCGTGGCGCGGAGCAGCGCGCGCGCAGCACTCTCGTGTGCTGAGCTTT 6217
Db 464 PheSerPheSerValProAlaGlyGlnProArgProSerHisHisGlyValPheAlaPhe 483
QY 6218 CTGGTGACCCATCCCTTACAGCTTGTGTGTCGCCCGC 6259
Db 484 LeuValThrProSerProTyrGluLeuCysAlaValProArg 497

RESULT 3
S37284
cytochrome P450 2D - bovine
N:Contains: Oxidoreductase (EC 1.-.-.-)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 03-Mar-2000
C:Accession: S37284; S29295; S29862
R:Tsunekawa, Y.; Matsuo, Y.; Higuchi, R.; Ichikawa, Y.
Eur. J. Biochem. 208, 739-746, 1992
A:Title: Characterization of the cytochrome P-450IID subfamily in bovine liver. Nucleot
A:Reference number: S29295; PMID:93011103; PMID:1396678
A:Accession: S37284
A:Molecule type: mRNA
A:Residues: 1-500 <TSU>
A:Cross-references: EMBL:X68481; NID:G295; PIDN:CAA48501.1; PID:G296
A:Experimental source: clone pBVL 180
A:Accession: S29295
A:Molecule type: mRNA
A:Residues: 14-111,'R',113-131,'R',133-162,'L',164-178,'G',180-219,'F',221-247,'R',249-2
A:Cross-references: EMBL:X68013; NID:G293; PIDN:CAA48149.1; PID:G294
A:Experimental source: clone pBVL 76
C:Genetics:
A:Gene: CYP2D
A:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F:305-468/Domain: cytochrome P450 homology <P45>
F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Alignment Scores:
Pred. No.: 2,94e-59 Length: 500
Score: 1318.50 Matches: 379
Percent Similarity: 30.36% Conservative: 47
Best Local Similarity: 27.01% Mismatches: 71
Query Match: 7.04% Indels: 906
Db: 1 Gaps: 12

```

```

US-09-820-788A-3 (1-10278) x S37284 (1-500)
QY 2078 ATGGGGCTA-----GAAGCACTGTGTGCTCCCTGCGCATGATAGTGGCCACTTCTCTG 2128
Db 1 MetGlyLeuLeuSerGlyAspThrLeuGlyProLeuAlaValAlaLeuLeuIlePheLeu 20
QY 2129 CTCCTGGTGGACCTGATGCACCGGCACCAACGCTGGGCTGCACCTACCCCGAGGTCCC 2188
Db 21 LeuLeuLeuAspLeuMetHisArgSerArgTrpAlaProArgTyrProProGlyPro 40
QY 2189 CTGCACCTGCCCGGCTGGCAACCTTCTGCATGTGACTTCCAGAACACACCATACTG 2248
Db 41 ThrProLeuProValLeuGlyAsn-LeuLeuGlnValAspPhe----- 54
QY 2249 CTTGCACAGGTGAGGAGAGGCTCTGGAGGGCGGAGAGTCTGAGGATGCGCCAC 2308
Db 55 -----GluAsp-Pro----- 57
QY 2309 ACCAGCAACATGGTGTGGTTAAACACAGGCTGGATCAGAAGCCAGGCTGAGAAG 2368
Db 57 ----- 57
QY 2369 GGAAGCAGGTTTGGGGACGTTCTCTGGGAAAGACATTTATACATGTCATGAAGGACTGG 2428
Db 57 ----- 57
QY 2429 ATTTTCCAAAGCCAAAGAGAGTAGGCAAGGCGCTGGAGGTGAGAGTGGACTTGGCAG 2488
Db 57 ----- 57
QY 2489 TGGCATGCAAGCCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACA 2548
Db 57 ----- 57
QY 2549 CCAGAAGGAAAGCGCTTGGGAATGGAAGATGAGTTAGTCTCTGAGTGGCGTTTAAATCAG 2608
Db 57 ----- 57
QY 2609 AAATCGAGGATGAAGGGGTGCTGACCCCGTTCAAACCTTTTGCACTGTGGGTCTCTCG 2668
Db 57 ----- 57
QY 2669 GGCCTCACTGCTCACCGGCATGACCATCATCTGGAATGGATGCTAACTGGGGCTCT 2728
Db 57 ----- 57
QY 2729 CGGCAATTTTGGTGACTCTTTCAGAGTCATACCTGGGTGACGATCCAAACTGAGTTCT 2788
Db 57 ----- 57
QY 2789 CCATCACAAGAGGTGTGACCCCAACCCCTGCCCCACAGATCAGGAGGTGGGTCTCTCTCT 2848
Db 57 ----- 57
QY 2849 TCCACCTGCTCACTCTCTGTTAGCCCCGGGGTCTGCAAGGTTCAAATAGGACTAGGACC 2908
Db 58 -----ArgProSerPheAsn----- 62
QY 2909 TGTAGTCTGGGGTGATCTCTGCTTTCACAAGAGGCGCTGACCCCTCCCTCTGCGAGTTGCGGC 2968
Db 63 -----GlnLeuArgA 66
QY 2969 GCCCTTTCGGGAGCGTGTTCAGCTTCAGCTGAGCTGGCGCTGGACCGCGGTGCTGCTCAATG 3028
Db 66 rgaArgPheGlyAsnValPheSerLeuGlnValTrpThrProValValLeuLeuAenG 86
QY 3029 GGCTGGCGCGCTGGCGGAGCGATGTGTGCCCGGAGGACACCGCCGAGCCGCGCCG 3088
Db 86 lylLeuAlaAlaValArgGluAlaLeuValTyrArgSerGlnAspThrAlaAspArgProp 106
QY 3089 CTGCGCCCATCTACAGGTCTCTGGGCTTCGGGCGCGCTTCCCAAGCAAGCGCGGTGGG 3148
Db 106 roProAlaValTyrGluHisLeuGlyTyrGlyProArg----- 118

```



```
QY 3149 GGACAGAGACCGCGTTTCGTGGGGCCCGGGTGGACAGTGACCGTAGCCCAAGCAGCGCC 3208
Db 118 ----- 118
QY 3209 GACAGGCGTGGGTCCTGGACGTGAAACAGAGATAAAGCCAGCAGTGGGCTGAGGAC 3268
Db 118 ----- 118
QY 3269 AGTGGSCCAGGAAACCACTGCACGGGGGAGGTGCGAGTCTGTGGGCTGGGAGGGGGCGG 3328
Db 118 ----- 118
QY 3329 GGCTACTGCCAGACCCGCCAGAACCCGCTGGGCGAGGCTGATGCTCGAAGTGCGCGGT 3388
Db 118 ----- 118
QY 3389 GGGGGGACCGGCCCTATCTCTGGGGCTCAGTGTGGGCGGAGCGGGCGGATCTTCCTTG 3448
Db 118 ----- 118
QY 3449 AGTGGAAAGGTGCTCAGGGTGGGACAGACGAGGTGGGGCCAAACCCCGCCCCAGGCAGG 3508
Db 118 ----- 118
QY 3509 GGAGCAATGTGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGACCCGGGCTAGGGACTG 3568
Db 118 ----- 118
QY 3569 CGGGAGACCTTGTGGAGCCACAGGTTGGAGTGGGTGGCGAGGGTGGGGCCCAAGCCCTT 3628
Db 118 ----- 118
QY 3629 CATGGCAACGCCACGTGTCCTGCCCGCCCGAGGGTGCATCTGTGCGCTATGGGCC 3688
Db 119 -----Ala-GluGlyValIleuAlaAlaArgTyrGlyAsp 129
QY 3689 GCGTGGCGCGAGCAGAGGGCTTCTCCGTGTCACCTTCGCAACTTGGGCTGGGCAAG 3748
Db 130 AlaTrpAlaGluGlnArgPheSerLeuThrLeuArgAsnPheGlyLeuGlyLys 149
QY 3749 AGTCGCTGAGCAGTGGGTGACCGAGAGGGCGCTGCTGTCGCGCTTCGCGGAC 3808
Db 150 LysSerLeuGluGlnTrpValThrGluGluAlaSerCysSerCysAlaAla 166
QY 3809 AAGCCGTTGATGGCGAGAGGGCACAAGCCGGAACCTGGAGAGGGCGGAGCGAG 3868
Db 166 ----- 166
QY 3869 AAGCAACCCCTTACCCGCATCTCCACCCCGAGGACGCCCTTTTCGCCCAACGGCCT 3928
Db 167 -----Phe-AlaAspGlnAlaGlyArgProPheSerProMetAspLeu 180
QY 3929 CTTGGCAAAAGCGGTGAGCAAGTGATCGCTCCCTCCCTCAGCTCGCGGCGCGCTTCAGTA 3988
Db 180 uLeuAsnLysAlaValSerAsnValIleAlaSerLeuThrPheGlyCysArgPheGluTy 200
QY 3989 CGACGACCTCGCTTCTCAGGCTGTGGACCTAGCTCAGGAGGAGTCAAGAGGAGTC 4048
Db 200 rAsnAspProArgIleIleLysLeuLeuAspLeuThrGluAspGlyLeuLysGluPr 220
QY 4049 GGGCTTCTCGCGAGGTGGGAGCGAGACCGAGGAGTCTCTGCGAGGGCGAGCTCTG 4108
Db 220 oAsnLeuValAlaGlyb 225
QY 4109 AGAGGTGCCGGGCTGGACTGGGCGCTCCGAAGGGCAGGATTTGCATAGATGGTTGGG 4168
Db 225 ----- 225
QY 4169 AAAGGACATTCACGAGAGACCCCACTGTAAAGAGGGCGCTGGAGGAGGGGACATCTCAG 4228
Db 225 ----- 225
QY 4229 ACATGGTCGTGGAGAGAGTGTGCCCGGGTCAGGGGGCACCGAGAGAGGCCAAAGGACTCTG 4288
Db 225 ----- 225
QY 4289 TACCCCCCTCCAGCTTGGAGATTTTCGATTTTAGTTTCTCCTCTGGGCAAGGAGAGGG 4348
Db 225 ----- 225
QY 4349 TGGAGGCTGCACCTTGGGGGAGGACTTGTGTGAGGTCAGTGTAAAGCAGCAGGAGGCCCTG 4408
Db 225 ----- 225
QY 4409 GGTCTACTGGAGATGGCTGGGGCCTGAGACTTGTCTCAGGTGAACGACAGCAGCAGGAGG 4468
Db 225 ----- 225
QY 4469 GATTGAGACCCCGTTCTGTCTGTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4528
Db 226 -----ValValGluAlaValProValLeuLeuSerI 236
QY 4529 CCCAGGCTGGCTGCAAGGCTCTACGCTTCCAAAGGCTTTCCTGACCCAGCTGGATGA 4588
Db 236 eProGlyLeuAlaAlaArgValPheProAlaGlnLysAlaPheMetAlaLeuIleAspG 256
QY 4589 GCTGTAAGTACGACAGAGTACCTGGGACCCAGCCAGCCACCCCGAGACTGACTGA 4648
Db 256 uLeuIleAlaGluGlnLysMetThrArgAspProThrGlnProProArgHisLeuThrAs 276
QY 4649 GGCCTTCTGGCAAGAGAGGAGAGAGTGTGAGTGTGCTGCCAGTGGGGGGCAAGGCTG 4708
Db 276 pAlaPheLeu -----AspGluValLys 283
QY 4709 TGGGTTGAACGTCCTCCAGGAGGAATGAGGGAGGCTGGGCAAAAGGTTGGACAGTGCATC 4768
Db 283 ----- 283
QY 4769 ACCCGCGAGCCGCATCTGGGCTGACAGGTGCAGAATTTGGAGGTCAATTTGGGGCTTACC 4828
Db 283 ----- 283
QY 4829 CGTTCTATCCCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGGAGGCCCTGAGAGCAG 4888
Db 284 -----GluAlaLysGlyAsnProGluSer 292
QY 4889 CTTCAATGATGAGAACCTGCGCATAGTGTGGGTAACTGTTCTTCCCGGATGTGTAC 4948
Db 292 rPheAsnAspGluAsnLeuArgLeuValValAlaAspLeuPheSerAlaGlyMetValTh 312
QY 4949 CACCTCGACCCAGCTGGGCTGCGCTGCTGCTCATGATCTTACACCTGGATGTGCGCG 5008
Db 312 rThrSerThrThrLeuAlaTrpAlaLeuLeuLeuMetIleLeuHisProAspValGln-- 331
QY 5009 TGAGCCAGCTGGGGCCCAAGCAGGGACTGAGGGAGGAGGTACAGCTGGGGGGCCCT 5068
Db 331 ----- 331
QY 5069 GGGCTTAGTGGGACACCCGGGGCTTCCAGCACAGCGGTGGCCAGGCTCTCTGTAGCCCTA 5128
Db 331 ----- 331
QY 5129 ACTTCTCCAAACACAGAGGAAGGAGAGTGTCCCCCTGGGTGCTGACCCCATTTGTGGGAGC 5188
Db 331 ----- 331
QY 5189 CATGCTGTCCAGTCCGTGTCCAAACAGGAGATCGACGCTGATAGGGCAGGTGCGGCGA 5248
Db 332 -----Arg--ArgValGlnGlnGluIleAspGluValIleGlyGlnValArgArg 347
QY 5249 CCAGAGATGGGTGACAGGCTCACATGCCCTACACCACTGCTGATTCACCGAGTTCAG 5308
Db 348 ProGluMetGlyAspGlnAlaLeuMetProPheThrValAlaValHisGluValGln 367
QY 5309 CGCTTTGGGGACATCATCCCCCTGAGTGTGACCCCATATGACATCCCGTGCATCGAAGTA 5368
```

Db 368 ArgPheAlaSerIleValProLeuGlyLeuProHisMetThrSerArgAspIleGluVal 387
Qy 5369 CAGGGCTTCGCATCCCTAAGGTAGGCTGGCGCCCTCTCACCCAGCTCAGCACCGC 5428
Db 388 GlnGlyPheHisIleProLys----- 394
Qy 5429 ACCTGGTGATAGCCCCAGCATGGCTACTGCCAGGTGGGCCCACTTAGAAACCTCGGCCA 5488
Db 394 ----- 394
Qy 5489 CCTAGTCCTCAATGCCACCACACTGACTGTCCCACTTGGGTGGGGGTCCAGAGTATAG 5548
Db 394 ----- 394
Qy 5549 GCAGGGCTGGCCCTGTCTCATCCAGAGCCCCGTCTAGTGGGGAGACAAACAGGACCTGCC 5608
Db 394 ----- 394
Qy 5609 AGAATGTTGAGGAGCCAGCGCCTGCAGGAGAGGGGGCAGTGTGGGTGCTCTGAGAGG 5668
Db 394 ----- 394
Qy 5669 TGTGACTGGCCCTGTGTGGGTTCGAGAGGTACTGTGGAGCTTCTGGGCGCAGGAC 5728
Db 394 ----- 394
Qy 5729 TAGTTGACAGATCCAGCTGTGTGCCAGGCAGTGTGTCCCGGTGTGTTGGTGGCAG 5788
Db 394 ----- 394
Qy 5789 GGGTCCAGCATCTTAGAGTCCAGTCCAGTCCCACTCTCACCTGCATCTCTGCCCCAGGGAAC 5848
Db 395 -----Gly-Th 396
Qy 5849 GACTCATCACCAACCTGTCTATCGGTGTGAAGATGAGGCCGTCTGGGAGAAGCCCTT 5908
Db 396 rThrLeuIleThrAsnLeuSerValLeuLysAspGluThrValTrpGluLysProPh 416
Qy 5909 CCGCTTCCACCCGGAACACTCTCTGGATGCCAGGSCCACTTTGTGAAGCCGAGGCTT 5968
Db 416 eArgPheHisProGluHisPheLeuAspAlaGlnGlyArgPheValLysGlnGluAlaPh 436
Qy 5969 CCTGCCTTCTCAGCAGGTGCTGTGGGGAGCCGGCTCCCTGTCCCTCTCCGTGGAGTC 6028
Db 436 eIleProPhe-Ser----- 440
Qy 6029 TTGAGGGGTATCACCCAGGAGCAGGCTCACTGACGCCCTCCCTCCCTCCACAGCCCGC 6088
Db 441 -----AlaGlyArg 443
Qy 6089 CGTGCATGCTCGGGAGCCCTGGCCCGCATGGAGCTTCTCTCTCTTCACTCCCTCG 6148
Db 444 ArgAlaCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrSerLeu 463
Qy 6149 CTGAGCAGCTTCAGCTTCTCCGTGGCGCGGAGAGCCCGGCCAGCCCACTCTCGTGC 6208
Db 464 LeuGlnHisPheSerPheSerValProAlaGlyGlnProArgProSerGluHisGlyVal 483
Qy 6209 GTCAGCTTCTGTGACCCCATCCCTCAGAGCTTTGTGCTGTGCCCGC 6259
Db 484 PheAlaPheLeuValThrProAlaProTyrGlnLeuCysAlaValProArg 500
RESULT 4
JE0258
cytochrome P450 2D23 - rabbit
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Mar-2000
C:Accession: JE0258
R:Yamanoto, Y.; Ishizuka, M.; Takada, A.; Fujita, S.
J. Biochem. 124, 503-508, 1998
A:Title: Cloning, tissue distribution, and functional expression of two novel rabbit cy

A:Reference number: JE0258; MUID:98391821; PMID:9722658
A:Accession: JE0258
A:Molecule type: mRNA
A:Residues: 1-500 <YAM>
A:Cross-references: DDBJ:AB008784
A:Experimental source: liver
C:Comment: This protein shows high drug metabolizing activity.
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F:9-25/Domain: transmembrane #status predicted <TM1>
F:305-468/Domain: cytochrome CYP>
F:310-326/Domain: transmembrane #status predicted <TM2>
F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Alignment Scores:		
Pred. No.:	6,65e-59	Length:
Score:	1311.50	Matches:
Percent Similarity:	29.58%	Conservative:
Best Local Similarity:	27.01%	Mismatches:
Query Match:	7.00%	Indels:
DB:	1	Gaps:
		9

US-09-820-788A-3 (1-10278) x JE0258 (1-500)

QY	2078	ATGGGGCTA-----GAAGCACTGGTGGCCCTGGCCATGATAGTGGCATTCTTCCTG	2128
Db			
Db	1	MetGlyLeuLeuSerGlyGluAlaLeuAlaProLeuAlaValAlaLallePheLeu	20
QY	2129	CTCTGGTGAACCTGATGACCGGCCAACAGCTGGGTGCACGTCACCGCCAGGTCCT	2188
Db			
Db	21	LeuLeuValAspLeuMetHisLysArgProArgTIPAlaAlaArgTyrProProGlyPro	40
QY	2189	CTGCCACTCCCGGGCTGGCAACCTTGCTGCATGTGGACTTCCAGAACACACACATAC	2248
Db			
Db	41	ValGlyIleProGlyLeuGlyAsn-LeuLeuGlnValAspPheArgGlyIleProAsnCy	60
QY	2249	CTTCGACCAAGGTGAGGAGGAGGTCTCTGGAGGGCGCAGAGGTCTTGAGGATGCCCA	2308
Db			
Db	60	sPhe-Arg-----	62
QY	2309	ACCAGCAAAATGGTGTGGGTTAAACCAACAGCTGGATCAGAACCGGCTGAGAAG	2368
Db			
Db	62	-----	62
QY	2369	GGAAGCAGGTTTGGGGACGTTCTCTGGGAAGGACATTTATACATGCATGAAGGACT	2428
Db			
Db	62	-----	62
QY	2429	ATTTCCAAAGGCCAAGGAAGATAGGGCAAGGGCCTGGAGGTGGAGCTGGACTTGG	2488
Db			
Db	62	-----	62
QY	2489	TGGGCATGCAAGCCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTG	2548
Db			
Db	62	-----	62
QY	2549	CCAGAAGGAAGCCCTTGGGAATGGAAGATGAGTTAGTCTGAGTCCGCTTTAAATC	2608
Db			
Db	62	-----	62
QY	2609	AAATCAGGATGAAGGGGGTGCAGTGCACCGGTTCAAACCTTTTGCACTGTGGGT	2668
Db			
Db	62	-----	62
QY	2669	GGCCTCACTGCTACCGGCATGGACCATCATCTGGGAATGGGATGTAACTGGGGCT	2728
Db			
Db	62	-----	62
QY	2729	CGGCAATTTTGGTGACTCTTGCAAGTGCATACCTGGGTGACGCATCCAACTGAGT	2788
Db			
Db	62	-----	62
QY	2789	CCATCACAAGAGGTGTGACCCCAACCCCTGCGCCCAAGATCAGAGGCTGGGTCTCT	2848

RESULT 4
JEO258
cytochrome P450 2D23 - rabbit
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Mar-2000
C:Accession: JEO258
R:Yamamoto, Y.; Ishizuka, M.; Takada, A.; Fujita, S.
J. Biochem. 124, 503-508, 1998
A:Title: Cloning, tissue distribution, and functional expression of two novel rabbit cy

```
Db 62 ----- 62
QY 2849 TCCACCTGCTCACTCCTGCTAGCCCGGGTCTGTCCTCAAGGTTCAAATAGGACTAGGACC 2908
Db 62 ----- 62
QY 2909 TGTAGTCTGGGTGATCTCTGGCTTGACAAGAGCCCTGACCTCCCTCTGCAAGTTCGGCG 2968
Db 63 ----- 66
QY 2969 GCGCTTCGGGACGTGTTTACGCTTCAGCTTCAGCTGCGCTGGACGCGGTGCTGCTCAATG 3028
66 rGArgTyrGlyAspValPheSerLeuGlnLeuAlaTrpThrProValValLeuAsnG 86
QY 3029 GGTGGCGCGCTGCGCGAGGCGATGTCACCGCGGCGAGACACGCGCGACCGCGCCGCGC 3088
86 lYProAlaValIleArgGluAlaLeuValThrTyrGlyLeuAspThrAlaAspArgProP 106
QY 3089 CTGCGCCCATCTACCAAGTCTCTGGCTTCGGCGCGCTTCCCAAGGCAAGCGCGTGGG 3148
106 roAlaHisThrLeuGluProLeuGlyPheGlyProHis ----- 118
QY 3149 GGACAGACCGCGTTTCGTGGGCGCGGCTGGACAGTGCAGCTAGCCCAAGACGCGCC 3208
118 ----- 118
QY 3209 GACAGGCGCTGGGCTCTCGACGTGAACAAGATAAAGCCAGCGAGTGGCTGAGGAC 3268
118 ----- 118
QY 3269 AGTGGGCCAGAAACCACTGACACGGGGAGGTGCGAGTCTGTGGCTGGGAGGGGCGG 3328
118 ----- 118
QY 3329 GGTACTGCCACAGCCCGCAGAACGCGTGGCGAGGCTGATGCTCGAAGTGGCGGT 3388
118 ----- 118
QY 3389 GCGGGGACCGCCCTATGCTGGGGCTCAGTGTGGCGGGAGCGGGCGGATCTTCCTTG 3448
118 ----- 118
QY 3449 AGTGAAGAGTGTGTCAGGGTGGCGAGACGAGTGGGGCCAAACCCCGCCCGAGCAGG 3508
118 ----- 118
QY 3509 GGAGCAATGTGGTGAGCAAGAGTGGGCGCTGTGCCAGCTGGACCGGGCTAGGGAATG 3568
118 ----- 118
QY 3569 CGGGAGACCTTGTGGAGCCCGAGGTTGGAGTGGGTGGCGGAGGTGGGGCCCAAGGCCTT 3628
118 ----- 118
QY 3629 CATGGCAACCCACGTCGTCGCGCCCGCCAGGGGTGATCTGTCGCGCTATGGGCC 3688
119 ----- 129
QY 3689 GCGTGGCGGAGCAGAGCGCTTCTCCGTGTCCACTTCGCAACTTGGCGCTGGGCAAG 3748
130 AlaTrpArgGluGlnArgPheSerValSerThrLeuArgAsnPheGlyLeuGlyLys 149
QY 3749 AGTCCCTGAGCAGTGGGTGACCGAGAGGCGCGCTGCTTGTGCGCTTCGCGCTCGCGACC 3808
150 LysSerLeuGluGlnTrpValThrGluGluAlaThrCysLeuCysAlaAla 166
QY 3809 AAGCCGCTGGTGGGAGAGGACAAAGCGGGAAGTGGAGAGCGGGGAGCGGAG 3868
166 ----- 166
QY 3869 AAGGCAACCCCTTACCCGCACTCTCCACCCCGAGGACGCGCTTTCGCGCCCAAGGCGCT 3928
166 ----- 166
Db 167 -----
QY 3929 CTTTGACAAAGCGTGAGCAACGTCATCGCTCCCTCACCTCGGGCGCGCTTCAGTA 3988
180 uLeuAsnLysAlaValCysAsnValIleAlaSerLeuThrHisGlyCysArgPheGluTyr 200
QY 3989 CGACGACCCCTGCTTCTCAGCTCTGAGCCTAGCTCAGGAGGAGCTGCAAGAGGAGTTC 4048
200 rAspAspHisArgLeuThrArgLeuMetAspLeuThrGlnThrIleLeuLysGluSerTh 220
QY 4049 GGGCTTTCTGCGCAGGTGCGGAGCGAGACCGAGGAGTCTCTGACGGCGGAGCTCCTG 4108
220 rGlyAsnLeuProGln ----- 225
QY 4109 AGAGGTGCGGGGCTGGACTGGGGCTCCGAAGGCGAGATTTCATAGATGGTGGTGGG 4168
225 ----- 225
QY 4169 AAAGGACATTCAGGAGACCCACCTGTAAAGAGGCGCTGGAGGAGGAGGACATCTCAG 4228
225 ----- 225
QY 4229 ACATGCTCGTGGGAGAGGTGTCCCGGGTCAGGGGGCACAGGAGGCCAAGGACTCTG 4288
225 ----- 225
QY 4289 TACCCCGCTCCACGTTGGAGATTTCGATTTTAGTTTCTCTCTGGGCAAGGAGAGGG 4348
225 ----- 225
QY 4349 TGGAGGCTGGCATTGGGAGGAGCTTGGTGAGGTCACTGGTAAAGACAGGAGCGCCCTG 4408
225 ----- 225
QY 4409 GGTCTACCTGGAGATGGCTGGGGCTGAGACTTGTTCAGGTGAACGACAGACACAGGAG 4468
225 ----- 225
QY 4469 GATTGAGACCCGTTCTGCTGTGTAGTGTGAATGCTGTCCCGCTCTCTCTGCACAT 4528
226 ----- 236
QY 4529 CCCAGCGCTGGTGTCAAGGTCTTACGCTTCCAAAGGCTTCTTCCAGCCAGCTGATGA 4588
236 eProGlyLeuValAspLysValPheArgGlyGlnLysAlaPheMetAlaLeuLeuAspG 256
QY 4589 GGTGTAATGAGCAGAGTACCTGGGACCCAGCCAGCCAGCCAGGAGCTGACTGA 4648
256 uLeuValThrGluHisArgMetThrArgAspProAlaGlnProProArgAspLeuThrAs 276
QY 4649 GGCCTTCTGCGCAAGAGAGAGTGAAGTGGCTGCCACGCTGGGGGCAAGGCTGG 4708
276 pAlaPheLeuAspGlnValGluLys ----- 284
QY 4709 TGGGTGTAACCTGCCAGGAGGAATGAGGGGAGGCTGGGCAAAAGGTTGGACCACTGCATC 4768
284 ----- 284
QY 4769 ACCCGGAGCGCATCTGGGCTGACAGGTGAGAAATTGGAGGTCAATTTGGGGGGCTACCC 4828
284 ----- 284
QY 4829 CGTCTATCCCTGAGTATCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCAG 4888
285 ----- 292
QY 4889 CTTCAATGATGAGAACTCGCATAGTGTGGTAACTGTTCTTCCCTGGCGGATGGTAC 4948
292 rPheAsnAspAspAsnLeuArgLeuValValThrAspLeuPheAlaAlaGlyMetValTh 312
QY 4949 CACTCGACCACTGGCTGGGCTCTGCTCATGATCTTACCTGATGTGTGATGTGCGAGCG 5008
312 rThrSerIleThrLeuSerTrpAlaLeuLeuLeuMetIleLeuHisProAspValGln-- 331
```



```
Db      15  PheThrIleIlePheLeuLeuValAspLeuLeuMetHisArgGlnArgIleThrSer 34
QY      2171 CGTACCGCCAGGTCCCTCCGACCTGCGCGGCTGGCAACCTTCTGTCATGTGACTT 2230
Db      35  ArgTyrProGlyProValProTyrProValLeuGlyAsn-LeuLeuGlnIleAspPh 54
QY      2231 CAGAACACACCATACTGCTTCACCACGAGTGAGGAGGAGGTCTCGAGGCGGCGAGAG 2290
Db      54  eGlnAsnMetPro----- 58
QY      2291 TCCTGAGGATGCCCCACCACCAACATGGGTGGTAAACCAACAGGCTGGATCA 2350
Db      58  ----- 58
QY      2351 GAAGCCAGGCTGAGAAAGGAAGCAGGTTTGGGGGACGTTCTCTGGGGAAGGACATTATA 2410
Db      58  ----- 58
QY      2411 CATGGCATGAGACTGGATTTTCCAAAGGCCAAGGAAGTAGGCAAGGGCTGGAGG 2470
Db      59  -----AlaGlyPheGln-Lys----- 63
QY      2471 TGGAGCTGGACTTGGCAGTGGGATGCNAGCCCATTTGGGCAATATGTTATGGAGTACA 2530
Db      63  ----- 63
QY      2531 AAGTCCCTTCTGCTGACACCAAGAAAGGCGCTTGGGAATGAAGATGAGTTAGTCCTG 2590
Db      63  ----- 63
QY      2591 AGTGCCTTTAAATACGAATCGAGGATGAAGGGGTGCAGTGACCCGGTTCAAACCTT 2650
Db      63  ----- 63
QY      2651 TTGCACTGGGTCTCGGGCCTCACTGCTCACCAGCATGGACCATCATCTGGGAATGGG 2710
Db      63  ----- 63
QY      2711 ATGCTAACTGGGGCCTCTCGGCAATTTTGGTGACTCTTGCAAGGTCTACTCTGGGTGACG 2770
Db      63  ----- 63
QY      2771 CATCAAACCTAGTTCTCCATCACAGAAAGGTGTGACCCCACTGCTGCCACGATCAG 2830
Db      63  ----- 63
QY      2831 GAGGCTGGGTCTCCTCTCCACCTGCTCACTCTGCTAGTACCCCGGGGTGCTCCAAGT 2890
Db      63  ----- 63
QY      2891 TCAATAGGACTAGGACCTGTAGTCTGGGTGATCTCTGGTTGACAAGAGGCCCTGACCC 2950
Db      63  ----- 63
QY      2951 TCCCTCTCAGTTGCGCGCGCTTCGGGACGTTTTCAGCTGACGCTGGCTGGACGC 3010
Db      64  -----LeuArgCysArgPheGlyAspLeuPheSerLeuGlnLeuAlaPheGluS 80
QY      3011 CGGTGCTGCTCAATGGCTGGCGCTGGCGGACGTTTTCAGCTGACGCTGGCTGGACGC 3070
Db      80  erValValLeuAsnGlyLeuProAlaLeuArgGluAlaLeuValLysTyrSerGluA 100
QY      3071 ACAGGCGGACCGCGCTGCGCCGTGCGAGCGATGGTGAACCGCGCGGAGG 3130
Db      100 sphrAlaAspArgProProLeuHisPheAsnAspGlnSerGlyPheGlyProArgSer- 119
QY      3131 AAGGCAAGCGCGTGGGGGACAGACACCGCTTCTCGGGGCGCCGGGTGGACAGTGAC 3190
Db      120 Gln----- 120
QY      3191 CGTAGCCCAAGCAGCGCCGACAGGGCGTGGGGTCTTGGAGCTGGAAGAAACAGAGATAAAGGCC 3250
Db      120  ----- 120

3251  AGCGAGTGGGCTGAGGACAGTGGCGCAGGAAACACCTGCACGGGGAGGTGCGAGTCTG 3310
120  ----- 120
3311  TGGGCTGGAGGGGGGGGGGCTACTGCCCCAGACCCGCCAGAGCCCGGTGGGGGAGGCTG 3370
120  ----- 120
3371  ATGCGTGAAGTGGCGGTGGCGGGAGCGCGCTATGCTGCGGCTCAGTGTGGGCGGGA 3430
120  ----- 120
3431  CGGGCGGATCTTCTTGAGTGGAAAGGTGTCAGGGTGGGAGACGAGGTGGGGCCA 3490
120  ----- 120
3491  AACCCCGCCCAAGGAGGAGCAATGTGGGTGAGCAAGAGTGGGCCCTGTGCCCAGCT 3550
120  ----- 120
3551  GGACCGGCTAGGACTGCGGAGACCTTGTGGAGCGCCAGGGTTGGAGTGGGTGGCGGA 3610
120  ----- 120
3611  GGGTGGGCAAGGCTTCATGGCAACGCCACGTGTCGTCCGCCCGCAGGGGTGATC 3670
121  -----GlyValVal 123
3671  CTGTGCGCTATGGGCCCGCTGGCGCGAGAGAGCGCTTCTCCGTGTCCACTTGGCGC 3730
124  LeuAlaArgTyrGlyProAlaTrpArgGlnArgPheSerValSerThrPheArg 143
3731  AACTTGGGCTGGCAAGNAGTCGTGGAGAGTGGTGCACGAGAGAGCGCGCTGCCCT 3790
144  HisPheGlyLeuGlyLysSerLeuGluGlnTrpValThrGluAlaArgCysLeu 163
3791  TGTGCGCTTCGCCCAACGCGGTGGGTGTGGCAGAGGGCAAGAGCGGAACTG 3850
164  CysAlaAlaPheAla----- 168
3851  GGAAGCGGGGACGGAGAAAGCAACCTTACCCTGATCTCCCAACCCCGAGGACGCC 3910
169  -----AspHis-Ser-----GlyPhePr 174
3911  CTTTGGCCCAAGCGCTCTTGGACAAAGCCGTGAGCAACGTGATCGCTCCCTCACCTG 3970
174  oPheSerProAsnThrLeuLeuAspLysAlaValCysAsnValIleAlaSerLeuLeuPh 194
3971  CGGGCGCGCTTCGAGTACGACGCCCTCGCTTCTCAGGCTGTGAGGACCTAGCTCAGGA 4030
194  eAlaCysArgPheGluTyrAsnAspProArgPheIleArgLeuLeuAspLeuLysAs 214
4031  GGGACTGAAGGAGTCTGGGCTTTCTGCGCGAGGTGGGAGGAGAGACCGAGGAGTCT 4090
214  pThrLeuGluGluGluSerGlyPheLeu----- 223
4091  CTCGAGGGCGAGCTCCTGAGAGGTGCGGGGTGAGTGGGCTCCGAGGGCAGGATT 4150
223  ----- 223
4151  TGCATAGATGGGTTTGGGAAAGGACATCCAGGAGACCCCACTGTAAAGAGGCGCTGGAG 4210
223  ----- 223
4211  GAGGAGGGACATCTCAGACATGCTGTGGGAGAGGTGTGCCGGGTTCAGGGGGCCACG 4270
223  ----- 223
4271  GAGAGGCCAAGGACTCTGTACCCCGTCCACGTTGGAGATTTTCGATTTTAGTTTCTCCT 4330
224  -----ProMet----- 225
```


A:Residues: 1-500 <POS>
A:Cross-references: GB:Y16417; NID:G2956687; PIDN:CAA76205.1; PID:G2956688
A:Accession: PC4502
A:Molecule type: protein
A:Residues: 2-57;249-273;408-430 <AXE>
A:Experimental source: liver
R:Axen, E.; Bergman, T.; Wikvall, K.
Biochem. J. 287, 725-731, 1992
A:Title: Purification and characterization of a vitamin D(3) 25-hydroxylase from pig liver
A:Reference number: S27177; MUID:93075023; PMID:1445236
A:Accession: S27177
A:Molecule type: protein
A:Residues: 2-17 <AX>
A:Experimental source: liver
R:Sono, H.; Sonoda, Y.; Sato, Y.
Biochim. Biophys. Acta 1078, 388-394, 1991
A:Title: Purification and characterization of cytochrome P-450(14DM) (lanosterol 14-alpha
A:Reference number: S17048; MUID:91316123; PMID:1859829
A:Accession: S17048
A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-11 <SON>
A:Note: 6-Leu was also found
C:Comment: This enzyme catalyzes the first step in the metabolic activation of vitamin D
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F:305-468/Domain: cytochrome P450 homology <P45>
F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Alignment Scores:
Pred. No.: 9.91e-57 Length: 500
Score: 1268.50 Matches: 370
Percent Similarity: 30.08% Conservative: 52
Best Local Similarity: 26.37% Mismatches: 75
Query Match: 6.77% Indels: 906
DB: 1 Gaps: 11

US-09-820-788A-3 (1-10278) x JC5819 (1-500)

QY 2078 ATGGGGCTA-----GAAGCAGTGGTCCCTCGCCATGATAGTGGCCATCTTCCG 2128
Db 1 MetGlyLeuLeuThrGlyAspLeuLeuGlyLeuLeuAlaMetValIlePheLeu 20
QY 2129 CTCTGTGTGACCTGATGACCGCCGACCAACGCTGGGGTGCAGCTACCGCCGAGGTCC 2188
Db 21 LeuLeuValAspLeuMetHisArgArgSerArgTyrAlaProArgTyrProGlyPro 40
QY 2189 CTGCCACTGCCCGGGCTGGGCAACCTTGTGTCATGTGGACTTCCAGAACACACCATCTG 2248
Db 41 MetProLeuProGlyLeuGlyAsn-LeuLeuGlnValAsnPheGln----- 55
QY 2249 CTTCCGACAGGTGAGGGAGGAGTCTCGGAGGGCGGACAGAGTCTTGAGATGCCCCACC 2308
Db 55 ----- 55
QY 2309 ACCAGCAACATGGGTGGTGGTGTAAACACAGGCTGGATCAGAGCCAGGTGAGAAGG 2368
Db 55 ----- 55
QY 2369 GGAAGCAGGTTTGGGGGACGTTCTCTGGGGAAGGACATTTATACATGGCATGAAGGACTGG 2428
Db 55 ----- 55
QY 2429 ATTTTCAAGGCCAAGGAAGAGTAGGGCAAGGGCCTGGAGGTGGAGCTGGACTTGGCAG 2488
Db 55 ----- 55
QY 2489 TGGGCATGCAAGCCCATTTGGGCAACATATGTTATGGAGTACAAAGTCCCTTCTGCTGACA 2548
Db 55 ----- 55
QY 2549 CCAGAAGGAAGGCCTTGGGAATGGAAGATGAGTTAGTCTCTGAGTCCGCTTTAAATCAGC 2608

Db 55 ----- 55
QY 2609 AAATCAGGATGAAGGGGTGCAGTGACCCCGGTTCAAACCTTTTGCACATGTGGTCTCTCG 2668
Db 55 ----- 55
QY 2669 GGCCTCACCTGCTACCGGCGATGGACCATCATCTGGGAATGGATGCTAACTGGGGCCTCT 2728
Db 55 ----- 55
QY 2729 CGGCAATTTGGTGACTCTTGAAGGTGATACCTGGGTGACGATCCAAACTGAGTTCT 2788
Db 55 ----- 55
QY 2789 CCATCACAAGGTGTGACCCCAACCCCTGCCCGCCAGATCAGGAGGTGGTCTCTCTCT 2848
Db 56 -----AspProArg-----Leu-SerP 61
QY 2849 TCCACCTGCTCACTCTCTGGTAGCCCGGGGGTCTGCAAGGTTCAAATAGGACTAGGACC 2908
Db 61 he----- 61
QY 2909 TGTAAGTCTGGGTGATCTCTGGCTTGCAAGAGGCCCTGACCCCTCTGACAGTTGCGGC 2968
Db 62 -----IleGlnLeuArg 66
QY 2969 GCGCTTCGGGACGCTGTTTACGCTGTCAGCTGGCCTGGAGCGCGGTGCTGCTCAATG 3028
Db 66 rgArgPheGlyAspValPheSerLeuGlnIleTyrArgProValValLeuAsnG 86
QY 3029 GCGTGGCGGCTGCGGAGGCGATGTGACCCCGGGGAGACACAGCCGCGACCGCCCGC 3088
Db 86 lyLeuAlaAlaValArgGluAlaLeuValSerHisSerHisGluThrSerAspArgPro 106
QY 3089 CTGCCCCCATCTACAGGCTCTGGGCTTCGGGCGCGTTCGCCAAGGCAAGCGCGGTGG 3148
Db 106 roValPheIleLeuGluHisLeuGlyTyrGlyProArgSer-Glu----- 120
QY 3149 GGACAGAGACCGCGTTTCCGTGGGCGCGGTGGACAGTGCAGTGCAGCCGTAGCCCAAGCAGCGC 3208
Db 120 ----- 120
QY 3209 GACAGGCGTGGGTCTCTGACGTAAGCAAGATAAAGGCCAGCGAGTGGGTGAGGAC 3268
Db 120 ----- 120
QY 3269 AGTGGGCCAGGAACCACTGCACGGGGAGGTGCGAGTCTGTGGCTGGAGGGGCGG 3328
Db 120 ----- 120
QY 3329 GGCTACTGCCAGACCCCGCCAGAAGCCCGTGGGCGAGGCTGATGCTCGAAGTGGCGGT 3388
Db 120 ----- 120
QY 3389 GCGGGGACCGCGCTATGCTCGGGCTCAGTGTGGGCGGAGCGGGCGGGATCTTCTCTTG 3448
Db 120 ----- 120
QY 3449 AGTGAAGGTGCTCAGGGTGGGCGAGAGAGAGTGGGGCCAAACCCCGCCAGGCGAG 3508
Db 120 ----- 120
QY 3509 GGAGCAATGTGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGCTAGGAGCTG 3568
Db 120 ----- 120
QY 3569 CGGAGACCTTGTGAGCGCCAGGGTTGAGTGGGTGGCGAGGTGGGGCCCAAGGCCCTT 3628
Db 120 ----- 120
QY 3629 CATGGCAAGCCCAAGTGTCCGTCCGCCCGCCCGAGGGTGATCTCTGTCGGCTATGGCCC 3688
Db 121 -----GlyValIleLeuAlaArgTyrGlyLys 129

3689 GGTGGCGGAGCAGAGGGCTTCTCCGCTGTCACCTTCGCGCAACTTGGGCGCTGGGCAAG 3748
Db |||||
130 AlaTPrpGluGlnArgPheSerValSerThrLeuArgAsnPheGlyLeuGlyLys 149
QY AAGTCGCTGGAGCAGTGGGTGACGAGGAGCGCGCTGCTCTTTGTGCGCTTCGCGCACC 3808
Db |||||
150 LysSerLeuGluGluTrpValThrGlnGluAlaSerCysLeuCysAlaAla 166
QY AAGCGGTGGGTGATGGGCAGAGGACAAAGCGGGAACCTGGGAAGCGGGGGAGGAG 3868
Db |||||
166 ----- 166
QY AAGCAACCCCTTACCGGCATCTCCACCCCGCAGAGCGCCCTTCGCCCAACGGCT 3928
Db |||||
167 -----Phe-AlaAspGlnAlaGlyArgProPheSerProAsnLe 180
QY CTTGACAAAGCGTGAGCAACGTGATCGCTCCCTCACCTCGCGCGCGCTTCAGTA 3988
Db |||||
180 uLeuAsnLysAlaValSerAsnValIleAlaSerLeuThrPheAlaArgArgPheGluTy 200
QY CGACGACCTCGCTTCCTCAGCTGCTGACCTAGCTCAGGAGGACTGAAGGAGGTC 4048
Db |||||
200 rAsnAspProArgMetLeuLysLeuLeuAspLeuValLeuGlyLeuLysGluGluVa 220
QY GGGCTTCTGCGGAGGTCCGAGCGGAGACCGAGAGTCTCTCGCGGCGAGCTCCTG 4108
Db |||||
220 lGlyLeuMetArgGln----- 225
QY AGAGTGGCGGGCTGGACTGGGCTCCGAGGGCAGGATTTGCATAGATGGGTTGGG 4168
Db ----- 225
4169 AAAGGACATTCCAGGAGACCCCACTGTAGAAGGGCCTGGAGGAGAGGGGACATCTCAG 4228
Db ----- 225
4229 ACATGTCGTGGAGAGGTGTCCCGGGTCAGGGGGCCACCAGGAGGCCAAGGACTCTG 4288
Db ----- 225
4289 TACCCCGCTCCAGTTGGAGATTTCGATTTTAGTTTCTCTCTGGCAAGGAGAGAGG 4348
Db ----- 225
4349 TGGAGGCTGGCACTTGGGAGGGAGCTTGGTGAGGTCACTGGTGAAGGACAGGCGCCCTG 4408
Db ----- 225
4409 GGTCTACCTGGAGATGGCTGGGGCTGAGACTTGTCCAGGTGAACGCAGAGCACAGGAGG 4468
Db ----- 225
4469 GATTGAGACCCGTTCTGTCTGTGTAGTGTGAATGCTGTCGCCGCTCCTCTGCACAT 4528
Db |||||
226 -----ValLeuGluAlaMetProValLeuArgHisI 236
QY CCCAGCGCTGGTGGCAAGGTCTTACGTTCCAAAAGGCTTTCTGACCCAGCTGGATCA 4588
Db |||||
236 eProGlyLeuCysAlaLysLeuPheProArgGlnLysAlaPheLeuValMetIleAspG 256
QY GGTGTAAGTACGACAGATGACCTGGAGCCAGCCCGCAGCCAGCCGAGACCTGACTGA 4648
Db |||||
256 uLeuIleThrGluHisLysMetThrArgAspLeuAlaGlnProProArgAspLeuThrAs 276
QY GGCCTTCTGGCAAGAGAGGAGAGTGTGAGTGTGCTGCCAGGTGGGGGCAAGGGTGG 4708
Db |||||
276 pAlaPheLeuAspGluMetLys----- 283
QY TGGGTTGAACGTCACGAGGAGGAATGAGGGGAGGCTGGGCAAAAGGTTGGACCAAGTGCATC 4768
Db |||||
283 ----- 283

4769 ACCGGCGAGCGGCATCTGGGCTGACAGGTGCAGAAATTGGAGGTCAATTGGGGGCTACCC 4828
Db ----- 283
4829 CGTTCTATCCTCGATATCCTCTGGCCCTGCTCAGCCAAAGGGAGCCCTGAGAGCAG 4888
Db |||||
284 -----GluAlaLysGlyAsnProGluSerSe 292
QY CTTCAATGATGAGAACCTGCGCATAGTGGTGGGTAAACCTGTTCTCTTGGCGGATGGTAC 4948
Db |||||
292 rPheAsnAspGluAsnLeuArgLeuValAlaHisLeuPheSerAlaGlyMetIleTh 312
QY CACTCGACACGCTGGGCTGCGCTCCTCTCATGATCTCATACCTGGGATGTGCACGG 5008
Db |||||
312 rThrSerThrThrLeuAlaIlePheAlaLeuLeuLeuMetIleLeuHisProAspValGln- 331
QY TGAGCCAGCTGGGGCCCAAGCGAGGACTGAGGGAGGAGGTACAGCTGGGGGGCCCT 5068
Db ----- 331
5069 GGGCTTAGCTGGACACCCGGGGCTTCAGCACAGGCGTGGCCAGGCTCTGTGAAGCCTA 5128
Db ----- 331
5129 ACTTCTCCAAACACAGAGGAGGAGAGTGTCCCTGGGTGCTGACCCCATTTGTGGGAGC 5188
Db ----- 331
5189 CATGTCCTCCAGTCCGTGTCACAGGAGATCGACGACGTGATAGGACAGGTGGCGGA 5248
Db |||||
332 -----Arg-ArgValGlnGlnGluIleAspGluValIleGlyHisValArgGln 347
QY CCAGAGATGGGTGACAGGCTCACATGCCCTACACCACTGCGGTGATTCAGAGGTGACG 5308
Db |||||
348 ProGluIleLysAspGlnAlaLeuMetProPheThrLeuAlaValLeuHisGluValGln 367
QY CGCTTTGGGACATCATCCCTGAGTGTGACCATATGACATCCGTCGATCAAGTCA 5368
Db |||||
368 ArgPheGlyAspIleValProLeuGlyValAlaHisMetThrSerCysAspIleGluVal 387
QY CAGGCTTCCGATCCCTAAGTAGGCGCTGGCGCCCTCTCTCACCCCACTCAGCACCCAGC 5428
Db |||||
388 GlnGlyPheLeuIleProLys----- 394
QY ACCTGGTGAAGCCCCAGCATGGCTACTGCGCAGGTGGGCCCACTCTAGGAACCCCTGGCCA 5488
Db ----- 394
5489 CCTAGTCTCAATGCCACCACACTGACTGTCTCCCACTTGGGTGGGGGTCCAGAGTATAG 5548
Db ----- 394
5549 GCAGGCTGGCTGTTCATCCAGAGCCCCCTGTAGTGGGGAGAGAGGGGCACTGTGGTGCCTCTGAGAGG 5608
Db ----- 394
5609 AGAATGTTGAGAGGACCCAGCGCTGCAGGGAGAGGGGGCACTGTGGTGCCTCTGAGAGG 5668
Db ----- 394
5669 TGTGACTGGCGCTGCTGTGGGGTCCGAGAGGGGTACTGTGGAGCTTCTCGGGCGCAGGAC 5728
Db ----- 394
5729 TAGTTGACAGAGTCCAGCTGTGTGCCAGGCACTGTGTCTCCCGTGTGTTTGTGGTGGCAG 5788
Db ----- 394
5789 GGGTCCCAGCATCTAGAGTCCAGTCCCACTCTCACCCCTGCATCTCTGCGCCAGGGAAC 5848
Db |||||
395 -----Gly-Th 396
QY GACACTCATCAACCACTGTTCATCGGTGCTGAAGGATGAGGCCGTCTGGGAGAGCCCTT 5908

2849 TCACCTGCTCACTCTGCTAGCCCGGGGTCGTCACAGGTTCAAATAGGACTAGGACC 2908
Db ----- 61
2909 TGTAGTCTGGGTGATCTCTGGCTTGACAGAGCCCTGACCTCCTCTGCACTTCCGCG 2968
Db -----Ser-GlnLeuGlnG 66
2969 GCGCTTGGGACGTGTTCACTGTCAGCTGACCTGAGCGCGGTGCTGCTCAATG 3028
Db -----
66 lYArgPheGlyAsnValPheSerLeuGluLeuAlaTrpThrProValValValLeuAsnG 86
3029 GCGTGGCGCCCTGCGCAGGCGATGTCACCGCGGCGAGACACGCGCAGCCCGCGC 3088
Db -----
86 lYLeuGluAlaValArgGluAlaLeuValHisArgSerGluAspThrAlaAspArgProp 106
3089 CTGCGCCCATCTACAGAGTCTCTGGCTTCGGCCCGCTTCCCAAGCCAGCGCGGTGG 3148
Db -----
106 roMetProIleTyrAspHisLeuGlyLeuGlyProGluSerGlnGly-Leu----- 122
3149 GCACAGAGACCGGTTTCGTCGGGCCCCGGGTGGACAGTACCGTAGCCCAAGCAGCGCC 3208
Db ----- 122
3209 GACAGGGCTGGGTCCTGGACGTGAACACAGAGATAAAGCCAGCGAGTGGGCTGAGGAC 3268
Db ----- 122
3269 AGTGGGCCAGGAAACACCTGCACGGGGAGGTGCGAGTCTGTGGGCTGGAGGGGGCGG 3328
Db ----- 122
3329 GGCTACTGCCACAGCCCGCCAGAACCCCGTGGCGAGGCTGATCGTCGAAGTGGCGGT 3388
Db ----- 122
3389 GCGGGGACCGCCTATGCTGCGGGCTCAGTGTGGGGGACGGCGGATCTTCTCTTG 3448
Db ----- 122
3449 AGTGAAGAGTGTTCAGGTGGGCGAGACGAGGTGGGGCCAAACCCCGCCAGGCAGG 3508
Db ----- 122
3509 GGAGCAATGTGGGTGAGCAAGAGTGGCCCTGTGCCAGCTGGACCGGGCTAGGGACTG 3568
Db ----- 122
3569 CGGGAGACCTTGTGGAGCGCCAGGTTGGAGTGGGTGGCGAGGTTGGGGCCAAAGCCCTT 3628
Db ----- 122
3629 CATGGCAACGCCAGTGTCTGTCGCCGCCCGCCAGGGGTGATCCTGTGCGCTATGGGCC 3688
Db -----PheLeuAlaArgTyrGlyArg 129
3689 GCGTGGCGGACAGAGCGCTTCTCGTGTCCACTTGGCGAACTTGGGCGCTGGGCAAG 3748
Db -----
130 AlaTrpArgGlnArgPheSerLeuSerThrLeuArgAsnPheGlyLeuGlyArg 149
3749 AAGTCGCTGGACGTGGGTGACCGAGGCGCGCTGCTTGTGCGCTTCGCGCTTCGCGACC 3808
Db -----
150 lYsSerLeuGluGlnTrpValThrGluAlaSerCysLeuCysAlaAla----- 166
3809 AAGCCGGTGGGTGATGGGCGAGAAGGGCACAAAGCGGGAACCTGGGAAGCGGGGACGGAG 3868
Db ----- 166
3869 AAGGCAACCCCTTACCGCATCTCCACACGCGCCCTTTCGCCCAAGCGGCT 3928
Db -----Phe-AlaGluGlnAlaGlyArgProPheGlyProGlyAlaLe 180
3929 CTTGGCAAAAGCGGTGAGCAACGTGATCGCTCCCTCACTTCGCGGGCGCGCTTCAGTA 3988

180 uLeuAsnLysAlaValSerAsnValIleSerSerLeuThrTyrGlyArgPheGluTyr 200
QY -----
3989 CGACGACCTTCGCTTCCTCAGGCTGCTGAGCTAGCTCAGGAGGAGCTGAAGGAGAGTC 4048
Db -----
200 rAspAspProArgLeuLeuGlnLeuLeuThrGlnGlnAlaLeuLysGlnAspSe 220
QY -----
4049 GGGCTTCTGCGCAGGTCGGAGGAGACCGAGGAGTCTCTGCAGGGCGAGCTCTG 4108
Db -----
220 rGlyPheLeuArgGlu----- 225
4109 AGAGTGGCGGGCTGGACTGGGGCTCCGAAGGSCAGATTGTGATAGATGGGTTTGGG 4168
Db ----- 225
4169 AAGAGCATTCAGGAGACCCACCTGTAAAGAGGCGCTGGAGGAGGAGGAGACATCTCAG 4228
Db ----- 225
4229 ACATGGTCTGGAGAGGTGTGCCCGGTGAGGGGGCACAGGAGAGCCCAAGGACTCTG 4288
Db ----- 225
4289 TACCCCGTCCACGTTGGAGATTTCGATTTTAGGTTTCTCTCTGGGCAAGGAGAGGG 4348
Db ----- 225
4349 TGGAGGCTGGCACTTGGGAGGAGCTTGGTGAGGTCAGTGTGTAAGGACAGCAGGCCCTG 4408
Db ----- 225
4409 GGTCTACCTGGAGATGGCTGGGGCTGAGACTTGTCCAGGTGAACGCAGAGCACAGGAGG 4468
Db ----- 225
4469 GATTGAGACCCGTTCTGCTGTGTAGTGTGAACTGCTGAACTGCTCCCGCTCTCTGCACAT 4528
Db -----AlaLeuAsnSerIleProValLeuLeuHisI 236
4529 CCCAGCGCTGCTGGCAAGTCTTACGCTTCCAAAGGCTTTCCTGACCCAGCTGGATGA 4588
Db -----
236 eProGlyLeuAlaSerLysValPheSerAlaGlnLysAlaIleIleThrLeuThrAsnG 256
QY -----
4589 CTGTGTAACCTAGCACAGGATGACCTGGGACCCAGCCAGCCACCCGAGACCTGACTGA 4648
Db -----
256 uMetIleGlnGluHisArgLysThrArgAspProThrGlnProProArgHisLeuIleAs 276
QY -----
4649 GGCCTTCTGCGCAAGAGAGAGAGTGGAGTGGTCCACCGTGGGGGCGCAAGGGTGG 4708
Db ----- 284
4709 TGGGTTGAACGTCCTCCAGAGGAATGAGGGGAGGCTGGGCAAAAGGTTTGGACAGTGCATC 4768
Db ----- 284
4769 ACCCGGAGCGCGCATCTGGGCTGACAGGTGCAGAAATTGGAGGTCAATTGGGGGCTACCC 4828
Db ----- 284
4829 CGTTCTATCCCTGAGTATCCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCAG 4888
Db -----AlaLysGlyAsnProLysThrSe 292
4889 CTTCAATGATGAGAACCTGCGCATAGTGTGGGTAACTGTTCTCTGCGGGATGGTGAC 4948
Db -----
292 rPheAsnGluGluAsnLeuCysMetValThrSerAspLeuPheIleAlaGlyMetValSe 312
QY -----
4949 CACCTGACACGCTGGCGCTGGGGCTCTCTCATGATCTACACCTGGATGTGCAGCG 5008
Db -----
312 rThrSerIleThrLeuThrTrpAlaLeuLeuLeuMetIleLeuHisProAspValGln-- 331
QY -----
5009 TGAGCCACAGCTGGGGCCCAAGGCGAGGACTGAGGGAGGAGGGGTACAGCTGGGGGCCCT 5068

Db 331 ----- 331
QY 5069 GGGCTTAGCTGGGACACCCGGGGCTTCCAGACAGGCGTGGCCAGGCTCCTGTAGCGCTA 5128
Db 331 ----- 331
QY 5129 ACTTCTCCAAACACAGGAGGAGAGTGTCCCTGGGTGCTGACCCATTGTGGGGACG 5188
Db 331 ----- 331
QY 5189 CATGTCTGTCCAGTCCGTGTCCAAACAGGAGATCGACAGCTGATAGGCGAGTGGCGCGA 5248
Db 332 ----- Arg-ArgValGlnGlnGlnLeu 347
QY 5249 CCAGAGATGGGTGACCAAGGTCTCATGCCCTACACCACTCCCGTGTATTCACGAGGTGCAG 5308
Db 348 ProGluMetGlyAspGlnThrArgMetProPheThrValAlaValIleHisGluValGln 367
QY 5309 CGCTTTGGGACATATCCCTGAGTGTGACCATATGACATCCCGTGTACATCGAAGTA 5368
Db 368 ArgPheGlyAspIleValProLeuGlyValProHisMetThrSerArgAspThrGluVal 387
QY 5369 CAGGGCTTCGGCATCCCTAAGTATAGGCTCGCGCCCTCTCACCCAGCTCACACACGAGC 5428
Db 388 GlnGlyPheLeuIleProlys ----- 394
QY 5429 ACCTGTGTATAGCCGCCAGCATGGCTACTGCCAGGTGGGGCCCACTCTAGGAACCTCGGCCA 5488
Db 394 ----- 394
QY 5489 CCTAGTCTCAATGCCACCACTGACTGTCCCACTTGGGTGGGGGTCCAGATATAG 5548
Db 394 ----- 394
QY 5549 GCAGGGCTGGCTGTCCATCCAGAGCCCCGTCTAGTGGGGAGACAAACAGGACCTGCC 5608
Db 394 ----- 394
QY 5609 AGAATGTTGGAGGACCCAGGCGCTGCAGGAGAGGGGCGAGTGTGGTGGCTCTGAGAGG 5668
Db 394 ----- 394
QY 5669 TGTGACTGGCCCTGTGTGGGTGCGAGAGGCTACTGTGAGCTTCTCGGGCGCAGGAC 5728
Db 394 ----- 394
QY 5729 TAGTTGACAGATCCAGCTGTGTGCCAGGAGTGTGTGTCCCGTGTGTGGTGGCAG 5788
Db 394 ----- 394
QY 5789 GGTGCCAGCATCTAGAGTCCAGTCCCACTCTCACCTGTGATCTCTGCCCGAGGGAAC 5848
Db 395 ----- Gly-Th 396
QY 5849 GACACTCATCAACCACTGTCTATCGGTGTGAAGATGAGCGCTGTGGGAGAGCCCTT 5908
Db 396 rThrLeuIleThrAsnLeuSerValLeuLysAspGluLysValTrpLysIleProPh 416
QY 5909 CGCTTCCACCCCGAACTTCTCTGGATGCCAGGCGCACTTTGTGAAGCCGGAGGCTTT 5968
Db 416 eArgPheTyrProGluHisPheLeuAspAlaGlnGlyHisPheValLysHisGluAlaPh 436
QY 5969 CTTGCTCTCTCAGAGGTGCTGTGGGGAGCCGGCTCCCTGTCCCTTCCCTGTGGAGTTC 6028
Db 436 eMetProPhe--Ser ----- 440
QY 6029 TTGACGGGTATCACCCAGGAGCAGGCTCACTGACGCCCCCTCCCTCCCAACAGGCGCG 6088
Db 441 ----- AlaGlyArg 443
QY 6089 CGTGCATGCTCGGGAGGCCCTTGGCCGCGATGAGCTCTTCTCTTCTTCACTCCCTG 6148
Db 444 ArgValCysLeuGlyGluProLeuAlaArgMetGluLeuPheLeuPheThrCysLeu 463

QY 6149 CTCACGACTTTCAGCTTCTCCTGGCCGCGGACACCCCGGCGCCAGCCACTCTCGTGC 6208
Db 464 LeuGlnArgPheSerValProAlaGlyGlnProArgProSerAspHisGlyVal 483
QY 6209 GTTCAGCTTTCGTGTACCCATCCCTAGAGACTTTGTGCTGTGCCCCGC 6259
Db 484 PheThrPheLeuLysValProAlaProPheGlnLeuCysValGluProArg 500
RESULT 9
S16872
cytochrome P450 2D3 - rat
N;Alternate names: cytochrome P450db3
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Mar-2000
C;Accession: S16872; E32970
R;Matsunaga, E.; Umeno, M.; Gonzalez, F.J.
J. Mol. Evol. 30, 155-169, 1990
A;Title: The rat P450 IID subfamily: complete sequences of four closely linked genes and 50 active site.
A;Reference number: S16871; MUID:90189185; PMID:2107330
A;Accession: S16872
A;Molecule type: DNA
A;Residues: 1-500 <MAT>
A;Cross-references: EMBL:X52028; NID:g57813; PIDN:CAA36270.1; PID:g57814
R;Matsunaga, E.; Zanger, U.M.; Hardwick, J.P.; Gelboin, H.V.; Meyer, U.A.; Gonzalez, F.
Biochemistry 28, 7349-7355, 1989
A;Title: The CYP2D gene subfamily: analysis of the molecular basis of the debrisoquine
A;Reference number: A32970; MUID:90057430; PMID:2819073
A;Accession: E32970
A;Molecule type: mRNA
A;Residues: 1-500 <MATS>
A;Cross-references: EMBL:J02868; NID:g203671; PIDN:AAA41002.1; PID:g203672
C;Genetics:
A;Gene: CYP2D3
A;Introns: 63/3; 121/1; 172/1; 225/3; 284/3; 332/1; 394/3; 442/1
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metal;
F;9-25/Domain: transmembrane #status predicted <TM1>
F;305-468/Domain: cytochrome P450 homology <CYP>
F;310-326/Domain: transmembrane #status predicted <TM2>
F;446/Binding site: heme iron (Cys) (axial ligand) #status predicted
Alignment Scores:
Pred. No.: 4,69e-52 Length: 500
Score: 1176.00 Matches: 348
Percent Similarity: 28.60% Conservative: 51
Best Local Similarity: 24.95% Mismatches: 93
Query Match: 6.28% Indels: 903
DB: 1 Gaps: 10
US-09-820-788A-3 (1-10278) x S16872 (1-500)
QY 2093 CTGCTGCCCTCGCCATGATAGTGGCCATCTCTCTGCTCCTGTGACCTGATGCACCG 2152
Db 9 LeuTrpProMetAlaIlePheThrValIlePheIleLeuLeuValAspLeuMethHisArg 28
QY 2153 CACCAACGCTGGGCTGCAGCTTACCGCCAGGTGCCCTGCCACTGCCGGCTGGGCAAC 2212
Db 29 ArgGlnArgTrpThrSerArgTyrProProGlyProValProTrpProValLeuGlyAsn 48
QY 2213 CTTGCTGCATGTGGACTTCCAGAACACACCATACTGCTTCGACGAGGTGAGGAGGAGGT 2272
Db 49 -LeuLeuGlnValAspLeuCysAsnMetProTyr-SerMet ----- 61
QY 2273 CTTGGAGGGCGCGACAGGTCTCTGAGGATGCCCCACCACCAACATCGGTGGTGGT 2332
Db 61 ----- 61
QY 2333 AAACACAGGCTGGATCAGAGCAGGCTGAGAGGGGAGCAGGTTTGGGGACGTTCC 2392
Db 61 ----- 61

QY 2393 TGGGAAGGACATTTATACATGCGATGAAGGACTGGATTTCCTCAAGGCCAAGGAGAGT 2452
Db 61 ----- 61
QY 2453 AGGCAAGGCCCTGGAGGTGGAGCTGGACTTGGCAGTGGGCATGCAAGCCCATTTGGCAA 2512
Db 61 ----- 61
QY 2513 CATATGTTATGAGTACAAAGTCCCTTCTGCTGACACCAAGAAAGGCCCTTGGGAATG 2572
Db 62 ----- 63
QY 2573 GAAGATGAGTTAGTCTCTGAGTCCGCTTTAAATCAGAAATCGAGATGAGGGGTGCAG 2632
Db 63 ----- 63
QY 2633 TGACCCGGTTCAAACTTTTGCACTGTGGGTCTCTGGGCTCACTGCTCACC GG CATGGA 2692
Db 63 ----- 63
QY 2693 CCATCATCTGGAATGGGATGTAATCTGGGGCTCTCGGCAATTTTGTGACTCTTGCAA 2752
Db 63 ----- 63
QY 2753 GGTATACCTGGGTGACGCATCCAACTGAGTTCTCTCCATCACAGAAGTGTGACCCCA 2812
Db 63 ----- 63
QY 2813 CCCCTGCCCCAGCATCAGGAGCTGGGTCTCTCTCTCCCTCTCCCTGCTCACTCTCTGTAGCC 2872
Db 63 ----- 63
QY 2873 CCGGGGTGTCCTCAAGTTCAATAGGACTAGGACCTGTAGTCTGGGGTGATCTCTGGTT 2932
Db 63 ----- 63
QY 2933 GACNAGGCCCTGACCCCTCCCTCTGCGATTGGGGCGCGCTTCGGGACGTGTTACGCC 2992
Db 64 ----- 74
QY 2993 TGCAGTGTGCTGGACGCGGTGTGTCTCAATGGGCTGGCGCGCTGGCGGAGGCGA 3052
Db 74 euGlnMetGlyTrpLysProValValIleAsnGlyLeuLysAlaValGlnGluLeuL 94
QY 3053 TGTGACCCGCGCGGAGGACACGCGCGACCGCCGCTGCGCCCATCTACCAAGGTCTCTGG 3112
Db 94 euValThrCysGlyGluAspThrAlaAspArgProGluMetProIlePheGlnHisIleG 114
QY 3113 GCTTCGGGCGGCTTCCCAAGCAAGCGCGGTGGGGGACAGACCGCGTTTCGCTGGG 3172
Db 114 lYTrGlyHisLysAla-Lys ----- 120
QY 3173 CCCCGGTGGACAGTGAACCGTAGCCCAAGCAGCGCGGACAGGGCGTGGGGTCTCGACGT 3232
Db 120 ----- 120
QY 3233 GAAACAGAGATAAAGGCCAGCGAGTGGGCTGAGGACAGTGGGCCAGGAAACCACTGCAC 3292
Db 120 ----- 120
QY 3293 GGGGAGGTGCGAGTCTGTGGCTGGAGGGGGGGGGCTACTGCCAGACCCGCCAGAA 3352
Db 120 ----- 120
QY 3353 GCCCGGTGGCGAGGCTGATGCTCGAAGTGGCGGTGGGGGACCGCGCTATGCTGCG 3412
Db 120 ----- 120
QY 3413 GGCTCAGTGTGGGCGGGACGGCGGGATCTTCTTGAAGTGAAGGTGCTCAGGGTGGGC 3472
Db 120 ----- 120

QY 3473 AGAGCAGAGTGGGGCCAAACCCCGCCAGCAGGGGAGCAAATGTGGTGTGAGCAAGAG 3532
Db 120 ----- 120
QY 3533 TGGGCCCTGTGCCAGCTGGACCGGGCTAGGACTTGGCGGAGACCTTGTGGAGCCCGAGG 3592
Db 120 ----- 120
QY 3593 GTTGAGTGGGTGGCGGAGGGTGGGGCCAAAGCCCTTCATGGCAACGCCACGCTGTCCGTC 3652
Db 120 ----- 120
QY 3653 CGGCCCCCAGGCGTGATCTGTGCGCTATGCGCCCGCGTGGCGGAGCAGAGCGCTTC 3712
Db 121 ----- 137
QY 3713 TCCGTGTCCACTTGGCGCACTTGGCGCTGGCGAAGAGTCTGGAGCAGTGGGTGAGCC 3772
Db 138 SerValSerThrLeuArgAsnPheGlyValGlyLysLysSerLeuGluGlnTrpValThr 157
QY 3773 GAGGAGCGCGCTGCTTTGTGCGCTTTCGCGACCAAGCCGGTGGGTGATGGGCGAGAAG 3832
Db 158 AspGluAlaSerHisLeuCysAspAla ----- 166
QY 3833 GGCACAAAGCGGAGCTGGGAGGGCGGGGACGGAGGCAACCCCTTACCCGCTATCTC 3892
Db 167 ---LeuThrAlaGluAlaGlyArg ----- 173
QY 3893 CCCACCCCGAGGCGCCCTTTCGCCCAACCGGCTCTTGACAAAGCCGTGAGCAACGT 3952
Db 174 -----ProLeuAspProTyThr-LeuLeuAsnLysAlaValCysAsnVa 188
QY 3953 GATCGCTCTCTACCTGCGGGCGCGCTTCGAGTACAGACCTGCTCTCTCAGGCT 4012
Db 188 lIleAlaSerLeuIleTyAlaArgPheAspTyArgLysProAspPheIleLysVa 208
QY 4013 GCTGACCTAGCTCAGGAGGAGCTGAAAGGAGAGTTCGGGCTTCTGCGGAGGTGCGGAG 4072
Db 208 lLeuLysIleLeuLysGluSerMetGlyGluGlnThrGlyLeuPheProGlu ----- 225
QY 4073 CGAGAGCCGAGGAGTCTCTGCGGGCGAGCTCTCGAGAGGTGCGGGGCTGGAGTGGG 4132
Db 225 ----- 225
QY 4133 CCTCGAAGGCGAGGATTTGTCATAGTGGGTTTGGGAAAGGACATTCAGAGAGACCCAC 4192
Db 225 ----- 225
QY 4193 TGTAAGNAGGCCCTGGAGGAGGGGACATCTCAGACATGCTGCTGGGAGAGGTGTGCC 4252
Db 225 ----- 225
QY 4253 CGGCTCAGGGGGCACAGAGAGGCCAAGGACTCTGTACCCCGCTCCACGTTGGAGATT 4312
Db 225 ----- 225
QY 4313 CGATTTTAGGTTTCTCTCTGGGCAAGAGAGAGGTGGAGGCTGGACATTGGGAGGGA 4372
Db 225 ----- 225
QY 4373 CTTGGTGGGTGAGTGGTAAGGACAGGAGGCCCTTGGGTCTACTGGAGATGGCTGGGGC 4432
Db 225 ----- 225
QY 4433 CTGAGACTTGTCCAGGTGAACGACAGAGCAGGAGGATTCAGACCCCGTTCTGTCTGGT 4492
Db 225 ----- 225
QY 4493 GTAGGTGCTGAATGCTGTCCCGCTCTCTCTGCAATCCACAGCGTGGTGGCAAGGTCT 4552
Db 226 ---ValIleAsnMetPheProValLeuLeuArgIleProGlyLeuAlaAspLysValPh 244
QY 4553 ACGCTTCCAAAGGCTTCTCTGACCCAGCTGGATGAGCTGCTAACTAGGACAGCATGAC 4612

A:Residues: 1-116,'D',118-345,'R',347-357,'F',359-406,'K',408-500 <ISH>
A:Cross-references: EMBL:M22330; NID:g203823; PIDN:AAA41049.1; PID:g203824
R:Macunaga E.; Umeno, M.; Gonzalez, F.J.
J. Mol. Evol. 30, 155-169, 1990
A:Title: The rat P450 IID subfamily: complete sequences of four closely linked genes and 50 active site.

A:Reference number: S16871; MUID:90189185; PMID:2107330

A:Accession: S16871

A:Molecule type: DNA

A:Residues: 1-345,'R',347-357,'F',359-406,'K',408-500 <MA2>

A:Cross-references: EMBL:X52027; NID:g57811; PIDN:CAA36269.1; PID:g57812

C:Genetics:

A:Gene: CYP2D2

A:Introns: 63/3; 121/1; 172/1; 225/3; 284/3; 332/1; 394/3; 442/1

C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology,
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F:9-25/Domain: transmembrane #status predicted <TM1>
F:305-468/Domain: cytochrome P450 homology <CYP>

F:310-326/Domain: transmembrane #status predicted <TM2>

F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Alignment Scores:

Pred. No.:	4,98e-52	Length:	500
Score:	1175.50	Matches:	347
Percent Similarity:	29.01%	Conservative:	60
Best Local Similarity:	24.73%	Mismatches:	90
Query Match:	6.28%	Indels:	906
DB:	1	Gaps:	11

US-09-820-788a-3 (1-10278) x B26822 (1-500)

QY	2078	ATGGGGCTA-----GAAGCACTGTGCCCCCTGGCCATGATAGTGGCCATCTTCCTG	2128
DB	1	MetGlyLeuLeuLeuGlyAspLeuTrpAlaValIlePheThrAlaIlePheLeu	20
QY	2129	CTCTGGTGGACCTGATGACCGGACCAACCGCTGGGCTGCACGCTACCCGCGAGTCCC	2188
DB	21	LeuLeuValAspLeuValHisArgHisPheTrpThrAlaHisTyProProGlyPro	40
QY	2189	CTGCACCTGCCGGCTGGGCACTTGTGATGTGACTTCCAGAACACACATCTG	2248
DB	41	ValProLeuProGlyLeuGlyAsn-LeuLeuGlnValAspPheGluAsnMetProTy-r	60
QY	2249	CTTCACCAAGTGGAGGAGGTCTTCGAGGGCGGACAGAGTCTGAGGATGCCCCACC	2308
DB	60	er-----	60
QY	2309	ACCAGAAACATGGGTGGTGGTTAAACACAGGCTGGATCAGAACGAGGCTGAGAAGG	2368
DB	60	-----	60
QY	2369	GGAAGCAGTTGGGGGACGTTCTCTGGGGAAGGACATTTATACATGATGAGACTGG	2428
DB	61	-----LeuTy-r-----	62
QY	2429	ATTTTCCAAAGCCAAGGAGTAGGAGGAGGCTGGAGGTGGAGCTTGGACTTGGCAG	2488
DB	62	-----	62
QY	2489	TGGGATGCAAGCCCATTTGGGCAACATATGTTATGGATGACAAAGTCCCTTCTGCTGACA	2548
DB	62	-----	62
QY	2549	CCAGAAGGAAGGCCCTTGGGAATGGAAGATGATGATGCTCTGAGTGGCCGTTAAATCAG	2608
DB	62	-----	62
QY	2609	AAATCAGGATGAAGGGGTGCAGTGACCCGGTTCAACCTTTTGCATGTGGGTCTCTCG	2668
DB	62	-----	62
QY	2669	GGCCTCACTGCTCACCGGATGGACCATCATCTGGGAATGGGATGCTAACTGGGGCCTCT	2728

Db	62	-----	62
QY	2729	CGGCAATTTTGGTGACTCTTCAAGGTATACCTGGGTGACGCATCCAACTGAGTTCT	2788
Db	62	-----	62
QY	2789	CCATCACAGAAGGTGTGACCCCTGACCCCTGACCCCTGACCCCTGACCCCTGACCCCT	2848
Db	62	-----	62
QY	2849	TCCACCTGCTCACTCTGTTAGTACCCCGGGGTCTGTCAGAGTTCAAAGTTCAAATAG	2908
Db	62	-----	62
QY	2909	TGTAGTCTGGGGTGATCTGCTGCTTGACNAGAGGCCCTGACCCCTCTGCTGAGTTG	2968
Db	63	-----LysLeuArgS	66
QY	2969	GCCGTTTCGGGGACGTGTTGAGCTGAGCTGGCTGGACCGCGGTGGTGTGCTCAATG	3028
Db	66	erArgTy-rGlyAspValPheSerLeuGlnIleAlaTrpLysProValValIleAsnG	86
QY	3029	GGTGGGGCGCTGCGAGCGGATGTGACCCCGCGGAGACACGCGCCGACCGCCGC	3088
Db	86	lyLeuLysAlaValArgGluLeuValThrTy-rGlyGluAspThrAlaAspArgProL	106
QY	3089	CTGCGCCCATCTACAGGTCTCGGGCTTCGGGCGCGCTTCCCAAGGCAAGCGCGGTGG	3148
Db	106	euLeuProIleTy-rAsnHisLeuGlyTy-rGlyAsnLysSer-Lys-----	120
QY	3149	GGACAGAGACCGCTTTCGTTGGGCGCGGTGGACGTGACCGTAGCCCAAGCAGCGCC	3208
Db	120	-----	120
QY	3209	GACAGGGCGTGGGTCTTGGACGTGAAACAGAGATAAAGCCAGCGAGTGGGTGAGGAC	3268
Db	120	-----	120
QY	3269	AGTGGGCCAGAAACACCTGCTCACGGGGAGTGGGAGTCTGTGGGTGGGAGGGGGCGG	3328
Db	120	-----	120
QY	3329	GGCTACTGCCACAGACCGCGGAGAGCCGCTGGGCGAGGCTGATCGCTCGAAGTGGCGGT	3388
Db	120	-----	120
QY	3389	GGCGGGACCGCGCTATGTCTGCGGGTCTAGTGTGGGCGGACGCGCGGAGTCTTCCTTG	3448
Db	120	-----	120
QY	3449	AGTGAAGGTGGTTCAGGGTGGGCGAGACAGAGTGGGGCCAAACCCCGCCAGGCAGG	3508
Db	120	-----	120
QY	3509	GGAGCAATGTGGGTGAGCAAGAGTGGGCCCTGTGCCAGCTGGACCGGGCTAGGGA	3568
Db	120	-----	120
QY	3569	CGGAGACCTTGTGGAGCGCCAGGGTTGAGTGGTGGCGGAGGTGGGGCCAAAGCCTT	3628
Db	120	-----	120
QY	3629	CATGGCAAGCCACGTGTGCTCCGCCCGGCGGTGATCTGCTGCGGTATGGGCC	3688
Db	121	-----GlyValValLeuAlaProTy-rGlyPro	129
QY	3689	CGCTGGCGCAGCAGAGCGCTTCTCGGTGTCCACCTTGGCGCAACTTGGGCTGGGCAAG	3748
Db	130	GluTrpArgGluGlnArgArgPheSerValSerThrLeuArgAspPheGlyValGlyLys	149
QY	3749	AAGTCGTGGAGCAGTGGGTGACCGAGAGCGCCCTGCTTGTGCGCTTCGCCCGACC	3808
Db	150	LysSerLeuGluGlnTrpValThrGluAlaGlyHisLeuCys-Asp-----	165

QY 3809 AAGCCGGTGGTGATGGGCAAGAGGACAAAGCGGAACTGGGAAGCGGGGACGGAG 3868
Db 166 -----Thr-PheAlaLysGluAlaGlu----- 172
QY 3869 AAGGCAACCCCTTACCGCATCTCCACCACCCAGGAGCCCTTTTCGCCCAACCGCCT 3928
Db 173 -----HisProPheAsnProSerIleLe 180
QY 3929 CTTGGACAAGCGGTGAGCAACGTGATCGCTCCCTCACCTGCGGGCGCGCTTCAGTA 3988
Db 180 uLeuSerLysAlaValSerAsnValIleAlaSerLeuValTyrAlaArgArgPheGluTyr 200
QY 3989 CGAGCACCTCTGCTTCTCAGGCTGGTGGACCTAGCTCAGAGGGGACTGAAGAGGAGTC 4048
Db 200 rGluAspProPhePheAsnArgMetLeuLysThrLeuLysGluSerPheGlyGluAspTh 220
QY 4049 GGGCTTTCTCGCGAGGTGGGAGCGAGACCGAGGAGTCTCTGCAGGGCGAGCTCCTG 4108
Db 220 rGlyPheMetAlaGlu----- 225
QY 4109 AGAGGTGCGGGGCTGACTGGGGCTCCGAAAGGCGAGGATTTGCATAGATGGTTGGG 4168
Db 225 ----- 225
QY 4169 AAAGGACATTCACAGAGACCCCACTGTAGAGGGGCTTGAGGAGGAGGGGACATCTCAG 4228
Db 225 ----- 225
QY 4229 ACATGTCGTGGAGAGGTGTCGCCGGTCAGGGGCAACGAGAGGCGAAGGACTCTG 4288
Db 225 ----- 225
QY 4289 TACCCCGTCCAGTTGGAGATTTCGATTTAGGTTTCTCCTCTGGCAAGGAGAGGG 4348
Db 225 ----- 225
QY 4349 TGGAGGCTGGCACTTGGGGAGGACTTGGTGAGTCAAGTGGTAAGCACAGGAGGCCCTG 4408
Db 225 ----- 225
QY 4409 GGTCTACTGGAGATGGCTGGGGCTTGAGACTTGTCCAGGTGAACGACAGCACAGGAG 4468
Db 225 ----- 225
QY 4469 GATTGAGACCCGTTCTGTCGTGTAGGTGTAATGTGTCTCCCGTCTCTCTGCACAT 4528
Db 226 -----ValLeuAsnAlaIleProIleLeuLeuGlnI 236
QY 4529 CCCAGCGCTGGCTGGCAAGTCTCTACGCTTCCAAAAGGCTTTCCTGCACCCAGCTGGATGA 4588
Db 236 eProGlyLeuProGlyLysValPheProLysLeuAsnSerPheIleAlaLeuValAspLys 256
QY 4589 GCTGCTAACTGACACAGGATGACCTGGGACCCAGCCAGCCACCCGAGACCTGACTGA 4648
Db 256 sMetLeuIleGluHisLysLysSerTrpAspProAlaGlnProProArgAspMetThrAs 276
QY 4649 GGCCTTCTCGGCAAGAGAGAGGTGAGTGGCTGCCACGCTGGGGGGCAAGGGTGG 4708
Db 276 pAlaPheLeuAlaGluMetGlnLys----- 284
QY 4709 TGGGTTGAACGTCCTCAGGAGGAATGAGGGGAGGCTGGGCANAAAGTTTGGACCAAGTGCATC 4768
Db 284 ----- 284
QY 4769 ACCCGGCGAGCCGCACTCGGGCTGACAGGTGCAGAAATTTGGAGGTCACTTTGGGGCTTACCC 4828
Db 284 ----- 284
QY 4829 CGTTCTATCCCTGAGTATCTCTCTCGGCCCTGCTCAGGCCAAGGGAGCCCTGAGAGCAG 4888
Db 285 -----AlaLysGlyAsnProGluSerSe 292

QY 4889 CTTCAATGATGAGAACCTCGCATAGTGTGGGTAACTGTTCTTCCTTGGCGGGGATGTGAC 4948
Db 292 rPheAsnAspGluAsnLeuArgLeuValIleAspLeuPheMetAlaGlyMetValTh 312
QY 4949 CACCTCGACACGCTGGCTGGGGCTCTCTCATGATCTACACTGGATGTGCGGCG 5008
Db 312 rThrSerThrThrLeuSerTrpAlaLeuLeuLeuMetIleLeuHisProAspValGln-- 331
QY 5009 TGAGCCACGCTGGGGCCCAAGCAGGAGACTGAGGGAGGAAGGTGTACAGCTGGGGGCCCT 5068
Db 331 ----- 331
QY 5069 GGGCTTAGCTGGGACACCGGGGCTTCCAGCACAGCGGTGGCCAGGCTCTGTAAAGCCTA 5128
Db 331 ----- 331
QY 5129 ACTTCTCTCAACACAGGAGGAAGAGAGTGTCCCTGGGTGTGACCCATTGTGGGGGACG 5188
Db 331 ----- 331
QY 5189 CATGTCTGTCTCAGTCCGTGTCCAACAGAGAGATCGACAGCTGTATAGGGCAGGTGCGGCGA 5248
Db 332 -----Arg--Arg-ValHisGluGluIleAspGluValIleGlyGlnValLeuArg 347
QY 5249 CCAGAGATGGGTGACCGGCTCACATGCCCTTACACACTGCCGTGATTCCAGAGGTGCAG 5308
Db 348 ProGluMetAlaAspGlnAlaArgMetProLeuThrAsnAlaValIleHisGluValGln 367
QY 5309 CGCTTTGGGACATCATCCCCCTGAGTGTGACCCATATGACATCCCGTGCACATCGAAGTA 5368
Db 368 ArgPheAlaAspIleValProThrAsnIleProHisMetThrSerArgAspIleLysPhe 387
QY 5369 CAGGCTCTCCGATCCCTAAGGTAGGCTGGCGCCCTCTCACCCAGCTCAGACCAGC 5428
Db 388 GlnGlyPheLeuIleProLys----- 394
QY 5429 ACCTGTGTATAGCCCCAGCATGGCTACTGCCAGTGGGGCCCACTCTAGGAACCCCTGGCCA 5488
Db 394 ----- 394
QY 5489 CCTAGTCTCAATGCCACCACTGACTGTCTCCCACTTGGGTGGGGGGTCCAGAGTATAG 5548
Db 394 ----- 394
QY 5549 GCAGGGCTGGCTGTCCATCCAGAGCCCCCGTCTAGTGGGGAGACAAACCAGGACCTGCC 5608
Db 394 ----- 394
QY 5609 AGAATGTTGGAGGACCCAGCGCTGCAGGGAGAGGGGGCAGTGTGGGTGCTCTGAGAGG 5668
Db 394 ----- 394
QY 5669 TGTGATGCGCCCTGCTGTGGGTTCGAGAGGGTACTGTGGAGCTTCTCGGGCGCAGGAC 5728
Db 394 ----- 394
QY 5729 TAGTTGACAGAGTCCAGCTGTGTGCCAGGCAGTGTGTCTCCCCCGTGTGTTTGGTGGCAG 5788
Db 394 ----- 394
QY 5789 GGGTCCCAGCATCTTAGAGTCCAGTCCCACTCTCACCTCGCATCTCTGCCCAGGGGAAC 5848
Db 395 -----Gly-Th 396
QY 5849 GACACTCATCAACACCTGTCTCATCGTGTGAAGGATGAGGCGCTGTGGAGAACGCCCTT 5908
Db 396 rThrLeuIleProAsnLeuSerSerValLeuGluAspGluThrValTrpGluLysProLe 416
QY 5909 CGCTTTCACCCCGAACACTTCTCTGATGCCAGGGCCACTTTGTGAAGCCGAGAGGCCCTT 5968
Db 416 uArgPheHisProGluHisPheLeuAspAlaGlnGlyAsnPheValLysHisGluAlaPh 436
QY 5969 CTTGCCCTTCTCAGCAGGTGCTGTGGGGAGCGCCGCTCTCCCTGTCCTCCCTGCGTGGAGTC 6028

Db 63 ----- 63
QY 2813 CCCTGCCCCACGATCAGGAGGCTGGGTCTCTCTTCCACCTGTCACTCTCTGGTAGCC 2872
Db 63 ----- 63
QY 2873 CCGGGGGTGTCCAAAGTTCAATAGACTAGACCTGTAGTCTGGGGTGATYCTGGCTT 2932
Db 63 ----- 63
QY 2933 GACAAGAGGCCCTGACCTCCCTCTGAGTTGGCGCGCTTTCGGGGACGTGTTTCAGCC 2992
Db 64 ----- 74
 LeuGlnHisArgTyrGlyAspValPheSerL 74
QY 2993 TGCAGCTGCGCTGACGCGCGGTGTCTCAATGGGCTGGCGCGCTGCGCGAGCGCA 3052
Db 74 euGlnMetGlyTrpLysProMetValIleValAsnArgLeuLysAlaValGlnGluValL 94
QY 3053 TGTGTACCGCGCGGAGGACACCGCGGACCGCGCGCTGGCGCCATCTACCAAGTCTCTGG 3112
Db 94 euValThrHisGlyGluAspThrAlaAspArgProProValProIlePheLysCysLeuG 114
QY 3113 GCTTCGGCGCGCTTCCCAAGGCAAGCGCGGTGGGGGACAGACCGCGTTTCCGTGGG 3172
Db 114 lyValLysProArgSer-Gln----- 120
QY 3173 CCGCGGGTGGACAGTACCGGTAGCCCAAGCAGCGCGCGCAGGGCGTGGGGTCTCTGGACGT 3232
Db 120 ----- 120
QY 3233 GAAACAGAGATAAGCCAGCGAGTGGGTGAGGACAGTGGGCGCAGGAAACCACTGAC 3292
Db 120 ----- 120
QY 3293 GGGGAGGTGCGAGTCTGTGGGTGGAGGGGGGGGCTACTGCCACACCGCGCAGAA 3352
Db 120 ----- 120
QY 3353 GCCCGGTGGCGGAGGCTGATGCTCGAAGTGGCGGTGGCGGGACCGCGCTATGTGGG 3412
Db 120 ----- 120
QY 3413 GGTCAAGTGGCGCGGACGCGCGGATCTTCTTGAATGGAAAGTGTCTAGGTGGGC 3472
Db 120 ----- 120
QY 3473 AGAGACGAGGTGGGCGCAACCCCGCCAGGCGGGGAGCAATGTGGGTGAGCAAGAG 3532
Db 120 ----- 120
QY 3533 TGGGCGCTGTGCCAGCTGGACCGGGCTAGGGACTCGCGGAGACCTTGTGGAGCGCAGG 3592
Db 120 ----- 120
QY 3593 GTTGGAGTGGGTGGCGGAGGTGGGGCCAAAGGCTTTCATGCGCAACGCCCTGTCCGTC 3652
Db 120 ----- 120
QY 3653 CCGCCCCCAGGGGTGATCTGTGCGCTATGGCCCGCGTGGCGGAGCAGAGGCGCTTC 3712
Db 121 ----- 137
 GlyValValPheAlaSerTyrGlyProGluTrpArgGluGlnArgPhe 137
QY 3713 TCGGTGTCCACCTTGGCAACTTGGGCTGGGCAAGTCTGCTGGAGCAGTGGGTGACC 3772
Db 138 SerValSerThrLeuArgThrPheGlyMetGlyLysLysSerLeuGluGluTrpValThr 157
QY 3773 GAGGAGCGCCCTGCTGCTTTGCTGCGCTTCCCGACCAAGCCGCTGGGTGTGATGGCAGAA 3832
Db 158 LysGluAlaGlyHisLeuCys-----Asp-AlaPheTh 168
QY 3833 GGCACAAAGCGGGAACCTGGGAAGCGGGGAGCGGAGAGGCAACCCCTTACCGGCATCTC 3892
Db 168 rAlaGlnAsn----- 171

QY 3893 CCCACCCCGCAGACGCCCCCTTTGCCCCCAACCGCCTCTTGGACAAAGCCGTGAGCAACGT 3952
Db 172 ----- 188
 GlyArgSerIleAsnProLysAlaMetLeuAsnLysAlaLeuCysAsnVa 188
QY 3953 GATCGCTCCCTCACTCGGGGGCGGCTTCAGTACGACGACCCCTCGCTTCTCTCAGGCT 4012
Db 188 lileAlaSerLeuIlePheAlaArgPheGluTyrGluAspProTyrLeuIleArgMe 208
QY 4013 GCTGACCTTAGCTCAGGAGGGACTGAAGGAGGAGTGGGGCTTCTGCGGAGGTGGGAG 4072
Db 208 tleuThrLeuValGluGluSerLeuIleGluValSerGlyPheIleProGlu----- 225
QY 4073 CGAGAGACCGAGAGTCTCTGCGAGGCGAGCTCTCTGAGAGGTGCGCGGGCTGAGTGGG 4132
Db 225 ----- 225
QY 4133 CCTCGAAGGGCAGGATTTGCAATAGATGGGTTTGGAAAGGACATTTCCAGGAGACCCAC 4192
Db 225 ----- 225
QY 4193 TGTAAGAGGGCTCGAGGAGGAGGGACATCTCAGACATGCTCGTGGGAGAGGTGTGCC 4252
Db 225 ----- 225
QY 4253 CCGGTACGGGGCACACGAGGAGGCCAAGACTCTGTACCCCGTCCACGTTGGAGATT 4312
Db 225 ----- 225
QY 4313 CGATTTTAGTTTCTCTCTGCGGCAAGGAGAGGGTGGAGCTGGCACCTTGGGAGGGA 4372
Db 225 ----- 225
QY 4373 CTTGGTGAAGTCAAGTAAAGGACAGGAGCGCCTGGGTCTACCTGGAGATGCTGGGC 4432
Db 225 ----- 225
QY 4433 CTGAGACTTGTCCAGGTGAACGAGAGCACAGGAGGATTGAGACCCGCTTCTGTCTGT 4492
Db 225 ----- 225
QY 4493 GTAGGTGTGAATGTGTCTCCCGTCTCTCTGCAATCCAGCGCTGGCTGGCAAGTCTCT 4552
Db 226 ----- 244
 ValLeuAsnThrPheProAlaLeuLeuArgIleProGlyLeuAlaAspLysValPh 244
QY 4553 ACGCTTCCAAAGGCTTCTCAGCCAGCTGATGAGTCTAACTGACACAGGATGAC 4612
Db 244 eGlnGlyGlnLysThrPheMetAlaPheLeuAspAsnLeuLeuAlaGluAsnArgThrTh 264
QY 4613 CTGGACCCAGCCAGCCACCCGAGACCTGACTGAGGCTTCTCTGGCAAGAGAGAA 4672
Db 264 rTrpAspProAlaGlnProProArgAsnLeuThrAspAlaPheLeuAlaGluValGluL 284
QY 4673 GGTGAGATGGTGTCCACCGTGGGGGGCAAGGGTGGTGGTTGAACGCTCCAGGAGGAAT 4732
Db 284 b----- 284
QY 4733 GAGGGAGGCTGGGCAAAAGGTTGGACCAGTGCATCACCCGGGAGCCGATCTGGGCTG 4792
Db 284 ----- 284
QY 4793 ACAGGTGCAGAAATTGGAGTCAATTTGGGGGCTACCCCGCTTCTATCCCTGAGTATCCTCT 4852
Db 284 ----- 284
QY 4853 CGGCGCTGTCAAGCCAAAGGAGCGCTGAGAGAGCTTCAATGATAGAACTGCGCAT 4912
Db 285 ----- 300
 AlaLysGlyAsnProGluSerPheAsnAspGluAsnLeuArgMe 300
QY 4913 AGTGGTGGTAACTTCTTCTTGGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4972
Db 300 tValValValAspLeuPheThrAlaGlyMetValThrThrAlaThrThrLeuThrTrpAl 320


```
QY 2153 CACCAACGCTGGCTGCACGCTACCCGCGCAGGTCCCTCGCACTGCCCGGCTGGGCAAC 2212
Db      |||||  ::|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
29 ArgGlnArgTrpThrSerArgTyrProProGlyProValProTrpProValLeuGlyAsn 48
QY 2213 CTTGCTGCATGTGGACTCCAGAACACACCATACTCTTCGACCAAGTGGAGGAGGT 2272
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
49 -LeuLeuGlnValAspLeuGlyAsnMetProTyr-Ser----- 60
QY 2273 CTTGGAGGCGGCAGAGGCTCTGAGGATGCCACCCACCAAGCAACATGGGTGGTGGTT 2332
Db      ----- 60
QY 2333 AAACACACGCTGGATCAGAGCCAGCTGAGNAGGGGAGCAGGTTTGGGGACGTTCC 2392
Db      ----- 60
QY 2393 TGGGGAAGGACATTATATACATGSCATGAAGGACTGGATTTCACAAAGGCCAAGGAAGGT 2452
Db      ::|||  -LeuTyr----- 62
QY 2453 AGGGCAAGGCGCTGGAGGTGGAGCTGGACTTGGAGTGGGCGATGCAAGCCCATTTGGSCAA 2512
Db      ----- 62
QY 2513 CATATGTTATGGAGTACAAAGTCCCTTCTGCTGACACCAAGAGAAAGCCCTTGGGAATG 2572
Db      ----- 62
QY 2573 GAAGATGAGTTAGTCTGAGTGCCTGTTAAATCACGAATCGAGGATGAAGGGGTGCAG 2632
Db      ----- 62
QY 2633 TGACCGGTTCAACCTTTTGCACTGTGGGTCTCGGGCTCACTGCTCACCGGCATGGA 2692
Db      ----- 62
QY 2693 CCATCATCTGGGAATGGGATGCTAACTGGGGCCTCTCGGCAATTTTGGTGACTCTTGCAA 2752
Db      ----- 62
QY 2753 GGTCAATACCTGGGTGAGCATCAAACTGAGTTCTCCATCACAGAGGTGTGACCCCA 2812
Db      ----- 62
QY 2813 CCCCTGCCCAAGATCAGGAGGTGGGTCTCTCTCTTCCACCTGCTCACTCTGGTAGCC 2872
Db      ----- 62
QY 2873 CCGGGGTGCTCAAGTTCAATAGGACTAGGACCTGTAGTCTGGGGTGATCTGGCTT 2932
Db      ----- 62
QY 2933 GACAAGAGCCCTGACCCCTCTCTGCTGCGGCGCGCTTCGGGAGCGTTTCAGCC 2992
Db      ::|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
63 -LysLeuGlnAsnArgTyrGlyAspValPheSerL 74
QY 2993 TGCAGCTGGCTGCAGCCGCTGCTCAATGGCTGGCGGCGCTGCGGAGGCGA 3052
Db      ::|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
74 euGlnMetAlaTrpLysProMetValValIleAsnGlyLeuLysAlaMetLysGluMetL 94
QY 3053 TGGTGACCCGCGGAGGACACGGCCGACCGCCGCTTCGCCCATCTACCAAGTCTCTGG 3112
Db      ::|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
94 euLeuThrCysGlyGluAspThrAlaAspArgProValProIlePheGluTyrLeuG 114
QY 3113 GCTTCGGGCGGCTTCCCAAGGCAGCGGGGTGGGGGACAGACCGGCTTTCCTGGG 3172
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
114 LysValLysProGlySer-Gln----- 120
QY 3173 CCCCGGGTGGACAGTGCAGTACCTGACCCCAAGCAGCGCGCAGAGGCGTGGGGTCTCGACGT 3232
Db      ----- 120
QY 3233 GAAACAGATAAAGCCAGCGAGTGGGCTGAGGACAGTGGGCGCAGGAAACCACTGCAC 3292
Db      ----- 120
Db 120 ----- 120
QY 3293 GGGGAGGTGCGAGTCTGTGGCTGGGAGGGGGCGGGCTACTGCCAGACCCCGCAGAA 3352
Db 120 ----- 120
QY 3353 GCCCGGTGGCGCAGGCTGATGCTGAAAGTGGCGGTGGCGGGACCGCGCCTATGCTGCG 3412
Db 120 ----- 120
QY 3413 GGTCTAGTCTGGCGGACGGCGGGATCTTCTTGTAGTGGAAAGTGTGTGAGGTGGGC 3472
Db 120 ----- 120
QY 3473 AGAGACAGGTGGGGCCAAACCCCGCCAGCAGGGAGCAATGTGGGTGAGCAAGAG 3532
Db 120 ----- 120
QY 3533 TGGGCCCTGTGCCAGCTGAGACCGGGCTAGGACTTGGCGAGACCTTGTGGAGCGCCAGG 3592
Db 120 ----- 120
QY 3593 GTTGGAGTGGGTGGCGAGGGTGGGGCCNAGGCCCTTCATGGCAAACGCCACCGTGTCCGTC 3652
Db 120 ----- 120
QY 3653 CCGCCCCCAGGGTGATCTGTGCGCTATGGGCCCGGTGGCGAGCAGAGCGCTTC 3712
Db 121 -----GlyValValLeuAlaProTyrGlyProGluTrpArgGluGlnArgPhe 137
QY 3713 TCCGTGTCCACCTTCGCAACTTGGGCTGGGCAAGAGTCTGCGAGCAGTGGGTGACC 3772
Db 138 SerValSerThrLeuArgAsnPheGlyLeuGlyLysLysLeuGluAspTrpValThr 157
QY 3773 GAGGAGGCGCTGCTTGTGCTGCGCTTCGCCGACCAAGCCCGTGGGTGATGGGCAAG 3832
Db 158 LysGluAlaAsnHisLeuCys-----Asp-AlaPheTh 168
QY 3833 GGCACAAAGCGGAACTGGGAAGCGGGGACGGAGAGGCAACCCCTTACCGCATCTC 3892
Db 168 rAlaGlnAla----- 171
QY 3893 CCCACCCCGCAGGACGCCCTTTCGCCCAACCGCTCTTGGACAAAGCCGTGAGCAACGT 3952
Db 172 -----GlyGlnProIleAsnProAsnProMetLeuAsnLysSerThrCysAsnVa 188
QY 3953 GATCGCTCCCTCACCTGCGGGCGCGCTTCGAGTACGACGACCCCTCGCTTCTCAGGCT 4012
Db 188 lIleAlaSerLeuIlePheAlaArgPheGluTyrGluAspProPheLeuIleArgMe 208
QY 4013 GCTGGACCTAGCTCAGGAGGACTGAAGGAGGTGCGGCTTTCGCGGAGTGGCGGAG 4072
Db 208 tLeuLysValLeuGluGlnSerLeuThrGluValSerGlyLeuIleProGlu----- 225
QY 4073 CGAGAGACCGAGGAGTCTCTGCGGGCGAGCTCTCTGAGAGGTGCGGGGCTGAGTGGG 4132
Db 225 ----- 225
QY 4133 CTTCCGAAGGGCAGGATTTCATAGATGGTTTGGGAAAGGACATTCACAGGAGACCCAC 4192
Db 225 ----- 225
QY 4193 TGTAGNAGGGCTCGAGGAGGGGAGACATCTCAGACATGCTCGTGGGAGAGGTGTGCC 4252
Db 225 ----- 225
QY 4253 CCGGTGAGGGGACACAGGAGGCGCAAGGACTCTGTACCCCGCTCCACGTTGGAGATT 4312
Db 225 ----- 225
QY 4313 CGATTTTAGTTTCTCTCTGGGCAAGGAGAGGGGTGGAGCTGGCACTTGGGGAGGGA 4372
```

Db 225 ----- 225
QY 4373 CTTGGTGGTCAAGTGAAGACAGGAGGCGCTGGGTCTACTGGAGATGGCTGGGGC 4432
Db 225 ----- 225
QY 4433 CTGAGACTTGTCCAGGTGAACGACAGACAGGAGGATTGAGACCCCGTTCTGTCTGT 4492
Db 225 ----- 225
QY 4493 GTAGTGTGAATGTCTCCGCTCTCTGTGACATCCAGGCGTGGCTGGCAAGTCTCT 4552
Db 226 -----ValLeuAsnAlaPheProIleLeuLeuArgIleProArgLeuAlaAspLysAlaLe 244
QY 4553 AGCTTCCAAAGGCTTCTCACCAGCTGATGAGCTGCTAACTGAGCAGCAGATGAC 4612
Db 244 uGlnGlyGlnLysSerPheIleAlaLeuLeuAspAsnLeuLeuThrGluAsnArgThrTh 264
QY 4613 CTGGGACCCAGCCAGCCACCCGAGACCTGACTGAGGCTTCTCTGGCAAGAGAGAA 4672
Db 264 rTrpAspProValGlnAlaProArgAsnLeuThrAspAlaPheLeuAlaGlnIleGlu 284
QY 4673 GGTGAGTGGTGTCCACCGTGGGGGCAAGGCTGTGGTGTGAACGTCCTCCAGAGGAAT 4732
Db 284 s ----- 284
QY 4733 GAGGGAGGCTGGGCAAAAGTTGGACAGTGCATCACCCGGCGAGCGCATCTGGGCTG 4792
Db 284 ----- 284
QY 4793 ACAGTGCAGAAATTGGAGTCAATTTGGGGCTACCCCGTTCTATCCCGTGTATCCTCT 4852
Db 284 ----- 284
QY 4853 CGGCCCTGTCTCAGGCCAAGGAGGAGCCCTGAGACAGCTTCAATGATGAGAACTCGCAT 4912
Db 285 -----AlaLysGlyAsnProGluSerSerPheAsnAspGluAsnLeuLeuMe 300
QY 4913 AGTGTGGTAACTTGTCTTCCGGGATGTGACCACTGACACCGTGGCGCTGGGG 4972
Db 300 tValValArgAspLeuPheGlyAlaGlyMetLeuThrThrSerThrLeuSerTrpAl 320
QY 4973 CTTCTGTCTGATGATCTACACTCGATGTGACGCTGAGCCAGTGGGGCCCAAGCA 5032
Db 320 alMetLeuMetIleLeuHisProAspValGln ----- 331
QY 5033 GGGACTGAGGAGGAAGGTTACAGCTGGGGGCCCTGGGGCTTAGCTGGGACACCGCGGC 5092
Db 331 ----- 331
QY 5093 TTCCAGCACAGGCTGGCCAGGCTCTGTAGCCTTAACCTTCTCCACACAGGAGGAAG 5152
Db 331 ----- 331
QY 5153 AGAGTGTCCCTGGTGTGTGACCCATTGTGGGACGCGATGTGTCCAGTCCGTGTCCAA 5212
Db 332 -----Arg--ArgValGln 335
QY 5213 CAGGAGATCGACGAGTATAGGGCAGGTGGCGACAGAGATGGGTGACAGGCTCAC 5272
Db 336 GlnGluIleAspGluValIleGlyGlnValArgHisProGluMetAlaAspGlnAlaHis 355
QY 5273 ATGCCCTACACACTCGCTGATTCACAGGTGAGCGCTTTGGGAGCATATCCCGCTG 5332
Db 356 MetProTyrThrAsnAlaValIleHisGluValGlnArgPheGlyAspIleValProVal 375
QY 5333 AGTGTGACCATATGACATCCCGTGCATCCAGTACAGGGCTCCGATCCCTAAGGTA 5392
Db 376 AsnLeuProArgIleThrSerHisAspIleGluValGlnAspPheLeuProLys --- 394
QY 5393 GGCCTGGCGCCCTCTCCACCCAGCTAGCACAGCAGCAGCTGTGTATAGCCCGCAGCATGGC 5452
Db 394 ----- 394

QY 5453 TACTCCAGAGTGGGGCCCACTCTAGGAACCCCTGGGCCACCTAGTCTCAATGCCACCACT 5512
Db 394 ----- 394
QY 5513 GACTGTCCCACTTGGGTGGGGGTTCAGAGTATATAGGAGGCGCTGGCCCTGTCCATCCAGA 5572
Db 394 ----- 394
QY 5573 GCCCCCGTCTAGTGGGGAGACAAACAGGACCTGCCAGAAATGTTGGAGGACCCAGCGCT 5632
Db 394 ----- 394
QY 5633 GCAGGAGAGGGGGCAGTGTGGTGCCTCTCAGAGGTGTGACTGGGCCCTGTGTGGGGT 5692
Db 394 ----- 394
QY 5693 CGGAGAGGTACTGTGGAGCTTCTCGGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTG 5752
Db 394 ----- 394
QY 5753 CCAGCAGTGTGTGTCCCGCTGTGTGTGGTGCGAGGGGTCCAGCATCTAGAGTCCAG 5812
Db 394 ----- 394
QY 5813 TCCCACTCTCACCTGTGATCTCTGCCCCAGGGAACGACATCATCAACCACTGTGTATC 5872
Db 395 -----Gly-ThrIleLeuLeuProAsnMetSerSe 404
QY 5873 GGTGTCAAGATGAGGCGCTCTGGGAGAGCCCTTCGCTTCCACCCCGCAACTTCTCT 5932
Db 404 rMetLeuLysAspGluSerValTrpGluLysProLeuArgPheHisProGluHisPheLe 424
QY 5933 GGATGCCAGGCGCACCTTTGTGAAGCCGAGGCGCTTCTGCTCTTCTCAGCAGGTGCCTG 5992
Db 424 uAspAlaGlnGlyHisPheValLysProGluAlaPheMetProPhe--Ser----- 440
QY 5993 TGGGAGCGCGGCTCCCTGTCTCCCTTCCCTGCTGGAGTCTTTCAGGGGTATCACCCAGGAGCC 6052
Db 440 ----- 440
QY 6053 AGGCTCACTGACGCGCCCTCCCTCCACAGCGCGCTGCATGCTCGGGAGCCCTG 6112
Db 441 -----AlaGlyArgArgSerCysLeuGlyGluAlaLeu 451
QY 6113 GCCCGCATGGAGCTTCTCTCTTCACTTCCCTGCTGTCAGCACTTCTCAGCTTCTCGGTG 6172
Db 452 AlaArgMetGluLeuPheLeuPheThrCysLeuLeuGlnArgPheSerPheSerVal 471
QY 6173 GCCCGCGACAGCGCGCGCCAGCCTCTCGTGTGCTGCTGCTTCTGCTGACCCCATCC 6232
Db 472 ProAspGlyGlnProGlnProSerAsnSerGlyValTyrGlyIleLeuValAlaProSer 491
QY 6233 CCCTACGAGCTTGTGTGTGTCGCCCGCTAGAAATGGG 6268
Db 492 ProTyrGlnLeuCysAlaValAlaArgAspGlnGly 503

RESULT 14

149428

cytochrome P450 16a-ms2 - western wild mouse

N:Contains: oxidoreductase (EC 1.-.-.-)

C:Species: Mus spretus (western wild mouse)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Mar-2000

C/Accession: 149428

R:Suwayoshi, T.; Kobayashi, R.; Nishio, K.; Aida, K.; Moore, R.; Wada, T.; Handa, H.; Neg

Mol. Cell. Biol. 15, 4158-4166, 1995

A>Title: A nuclear factor (NF249) that binds to the male-specific P450 (Cyp 2d-9) gene

A/Reference number: A57454; MUID:95349581; PMID:7623810

A/Accession: 149428

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-504 <RES>

A/Cross-references: EMBL:U20088; NID:g951101; PIDN:AAC52246.1; PID:g951102

QY	2093	CTGGTGGCCCTGGCCATAGTAGTGGCCATCTTCTGCTCTCTGGTGAACCTGATGCAACCGG	2152
Db	9	LeuTrpProValAlaIlePheThrValIlePheIleLeuLeuValAspLeuThrHisGln	28
QY	2153	CACCAACGCTGGGCTGCACGCTACCCGCCAGGTCCCTGCACCTGCCTGGCGGTGGGCAAC	2212
Db	29	ArgGlnArgTrpThrSerArgTyrProGlyProValProTrpProValLeuGlyAsn	48
QY	2213	CTTGCTGCATGTGGAGTTCAGAACACACATATCTGCTCGACACAGGTAGGAGGAGGT	2272
Db	49	-LeuLeuGlnValAspLeuAspAsnMetProTyr-Ser-	60
QY	2273	CCTGGAGGGCGGCAGAGGTCTCGAGGTGCCACACACAGCAAAATGGGTGGTGGTT	2332
Db	60	-----	60
QY	2333	AAACACAGGCTGGATCAGAAAGCCAGGCTCAGAAAGGGAGCAGGTTTGGGGACGTTC	2392
Db	60	-----	60
QY	2393	TGGGGAAGGACATTTATACATGTCATGAAGGACTGGATTTTCCAAAGGCCAAGGAAGT	2452
Db	61	-----LeuTyr-	62
QY	2453	AGGCAAGGCCCTGGAGGTGGAGCTGGACTTGGCAGTGGGCATGCAAGCCCATTTGGCAA	2512
Db	62	-----	62
QY	2513	CATATGTTATGGAGTACAAAGTCCCTTCTGCTGACACACAGAGAAAGGCTTTGGGAATG	2572
Db	62	-----	62
QY	2573	GAAGATGAGTTAGTCTGAGTGCCGTTTAAATCAGAAATCGAGGATGAAGGGGTGCAG	2632
Db	62	-----	62
QY	2633	TGACCCGGTTCAAACCTTTTGCACTGGGTCTCGGGCCTCAGTCTCACCGGCATGGA	2692
Db	62	-----	62
QY	2693	CCATCATCTGGGAATGGGATGCTAACTGGGGCCTCTCGGCAATTTTGGTGACTCTTGCAA	2752
Db	62	-----	62
QY	2753	GGTCATACCTGGGTGAGCGATCCAAACTGAGTTCTCCATCAGAGAAGGTGACCCCCA	2812
Db	62	-----	62
QY	2813	CCCCTGCCCAACGATCAGGAGGTGGGTCTCTCTCTTCCACCTGCTCACTCTCTGGTAGCC	2872
Db	62	-----	62
QY	2873	CCGGGGGTGCTCAAAGTTCAAAATAGGACTAGGACCTGTAGTCTGGGGTGATCCTGGCTT	2932
Db	62	-----	62
QY	2933	GACAAGGCCCTGACCCCTCCCTCTGAGTTGGCGCGCGCTTTCGGGGACGTTTCAGCC	2992

QY 4073 CGAGAGACGAGGAGTCTCTCAGGCGGAGCTCCTGAGAGGTGCGGGGCTGAGTGGGG 4132
Db 222 ----- 222
QY 4133 CCTCCGAGGCGAGGATTTGCATAGATGGTTTGGGAAAGGACATTCAGAGGACCCAC 4192
Db 223 ----- 225
QY 4193 TGTAAGAGGGCTGGAGAGGAGGGACATCTCAGACATGCTGCGGAGAGGTGGTGGC 4252
Db 225 ----- 225
QY 4253 CGGGTCAGGGGACACGAGAGGCCAAGACTCTGTATCCCGCTCCAGTTGGAGATT 4312
Db 225 ----- 225
QY 4313 CGATTTTAGGTTTCTCTCTGGGCAAGAGAGGGTGGAGCTGGCACTTGGGAGGGA 4372
Db 225 ----- 225
QY 4373 CTTGGTGGGTCAAGTAAAGACAGGAGGCCCTGGGTCTTACCTGGAGATGGTGGGC 4432
Db 225 ----- 225
QY 4433 CTGAGACTTGTCCAGGTGAACGACAGACAGAGGGATTGAGACCCCGTTCTGTCTGT 4492
Db 225 ----- 225
QY 4493 GTAGGTGCTGAATGTCTCCCGTCTCTCGACATCCAGCGCTGGTGGCGAGGTCT 4552
Db 226 ----- 244
QY 4553 ACCTTCCAAAGGTTTCTCAGCCAGCTGATGAGTGTCTAATGATGAGACAGGATGAC 4612
Db 244 eGlnGlyGlnLysSerPheMetAlaIleLeuAspAsnLeuThrGluAsnArgThrTh 264
QY 4613 CTGGGACCCAGCCAGCCAGACCTGATGAGTGTCTGAGCTTCTGGCAAGAGGAA 4672
Db 264 rTpaAspProAspGlnProArgAsnLeuThrAspAlaPheLeuAlaGluIleGlu 284
QY 4673 GGTGAGAGTGGTGCACCGTGGGGGCAAGGTGTGGTGTGAACGTCCAGGAGGAAT 4732
Db 284 s----- 284
QY 4733 GAGGGAGGCTGGGCAAAAGTTGACACAGTCATCACCCGCGAGCGCATCTGGGCTG 4792
Db 284 ----- 284
QY 4793 ACAGGTGCAGAAATTGGAGTCAATTTGGGGGCTACCCCGTTCTATCCCTGAGTATCCTCT 4852
Db 284 ----- 284
QY 4853 CGGCCCCTGCTCAGGCCAAGGAGCCCTGAGAGCAGCTTCAATGATGAGAACCTGGCAT 4912
Db 285 ----- 300
QY 4913 AGTGGTGGTAACTGTCTCTTGGGATGGTGGACCTCGACACGCTGGCTGGGG 4972
Db 300 tValValGlyAspLeuPheThrAlaGlyMetValThrThrSerThrLeuSerTpa 320
QY 4973 CTCTCTCATGATCTACACTGTGATGTGAGGTGAGCCAGCTGGGGGCCAAGGA 5032
Db 320 aLeuLeuMetIleLeuHisProAspValGln----- 331
QY 5033 GGGACTGAGGGAAGGTTACAGCTGGGGGCCCTTGGGCTTAGCTGGGACACCCGGGGC 5092
Db 331 ----- 331
QY 5093 TTCCAGCACAGGCGTGGCAGGCTCCTGTAAAGCTTAACCTTCTCTCAACAGGAGGAAG 5152
Db 331 ----- 331

QY 5153 AGAGTGTCCCTCCTGGTGTCTGACCCATTGTGGGAGCGCATGTCTGTCCAGTCCGTGTCAA 5212
Db 332 ----- 335
QY 5213 CAGAGATTCGACGACGTGATAGGCGAGGTGGCGCACAGAGATGGGTGACAGGCTCAC 5272
Db 336 GlnGluIleAspAlaValIleGlyGlnValArgHisProGluMetAlaAspGlnAlaHis 355
QY 5273 ATGCCCTACACACTCGCTGATTCACGAGTGCAGCGCTTTGGGACATCATCCCTCG 5332
Db 356 MetProTyrThrAsnAlaValIleGluValGlnArgPheGlyAspIleAlaProLeu 375
QY 5333 AGTGTGACCCATATGACATCCCGTACATCAAGTACAGGCTTCGCGCATCCCTAAGTA 5392
Db 376 AsnLeuProArgIleThrSerArgAspIleGluValGlnAspPheLeuIleProLys 394
QY 5393 GGCCTGGCGCTCTCTCACCCTCAGTCCAGCACCGACCTGGTGATAGCCCCAGCATGGC 5452
Db 394 ----- 394
QY 5453 TACTGCCAGGTGGGCCCACTCTAGGAACCTGGGCCCTAGTCTCTCAATGCCACACACT 5512
Db 394 ----- 394
QY 5513 GACTGTCCCACTTGGGTGGGGTCCAGAGTATAGGAGGGCTGGCCTGTCTCCATCCAGA 5572
Db 394 ----- 394
QY 5573 GCCCCCTGTAGTGGGGAGACAACCAGGACCTGCCAGAATGTTGGAGAGCCAGCGCT 5632
Db 394 ----- 394
QY 5633 GCAGGGAGAGGGGAGTGTGGTGGCTCTGAGAGGTGTGACTGGCGCTCTGTGGGGT 5692
Db 394 ----- 394
QY 5693 CGGAGAGGTACTGTGGAGCTTCTCGGCGCAGGACTAGTTGACAGAGTCCAGCTGTGTG 5752
Db 394 ----- 394
QY 5753 CCAGGACGTGTGTGTCCCTGTGTGTTGGTGGCAGGGGTCCAGCATCTAGAGTCCAG 5812
Db 394 ----- 394
QY 5813 TCCCCACTCTCACCTGTCATCTCTGCCAGGGAACGACATCATACCAACCTGTCTATC 5872
Db 395 ----- 404
QY 5873 GGTCTGAAGATGAGCGCTCTGGGAGAGCCCTTCGCTTCCACCCCGAACACTTCTCT 5932
Db 404 rValLeuLysAspGluThrValTrpGluLysProLeuHisPheHisProGluHisPheLe 424
QY 5933 GGATGCCAGGCGCACATTTGTGAAGCGGAGCGCTTCTGCTCTTCTCAGCAGGTGCCTG 5992
Db 424 uAspAlaGlnGlyHisPheValLysProGluAlaPheMetProPheSerAla----- 441
QY 5993 TGGGAGCCCGCT 6052
Db 442 ----- 443
QY 6053 AGGCTACTGAGCCCT 6112
Db 444 ----- 451
QY 6113 GCCCGCATGGAGCTTCT 6172
Db 452 AlaArgMetGluLeuPheLeuPheThrCysLeuLeuGlnArgPheSerIleSerVal 471
QY 6173 GCCCGCGGACAGCCCGGCGCAGCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 6232
Db 472 ProAspGlyGlnProGlnProSerAsnTyrArgValHisAlaIleProValAlaProPhe 491
QY 6233 CCCTACGAGCTTTGT 6259

```
Db 492 ProTyrGlnLeuCysAlaValMetArg 500
|||||:|||||:|||||:|||||
RESULT 15
A26822
debrisoquine 4-hydroxylase (EC 1.14.14.-) cytochrome P450 2D1 - rat
N:Alternative names: cytochrome P450 UT-7; cytochrome P450db1
C:Species: Rattus norvegicus (Norway rat)
C>Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 03-Mar-2000
C:Accession: A26822; A30495; E32970; C32970; A31579; J04158; S39761
R:Gonzalez, F. J.; Matsunaga, T.; Nagata, K.; Meyer, U.A.; Nebert, D.W.; Pastewka, J.; Ko
DNA 6, 149-161, 1987
A:Title: Debrisoquine 4-hydroxylase: characterization of a new P450 gene subfamily, regu
A:Reference number: A90957; MUID:87217961; PMID:3582092
A:Accession: A26822
A:Molecule type: mRNA
A:Residues: 1-504 <GN>
A:Cross-references: EMBL:M16554; NID:g203833; PID:AAA41054.1; PID:g203834
A:Accession: A30495
A:Molecule type: protein
R:Matsunaga, E.; Zanger, U.M.; Hardwick, J.P.; Gelboin, H.V.; Meyer, U.A.; Gonzalez, F. J
Biochemistry 28, 7349-7355, 1989
A:Title: The CYP2D gene subfamily: analysis of the molecular basis of the debrisoquine 4
A:Reference number: A32970; MUID:90057430; PMID:2819073
A:Accession: B32970
A:Molecule type: mRNA
A:Residues: 1-504 <MA1>
A:Cross-references: EMBL:J02867; NID:g203669; PID:AAA41001.1; PID:g203670
A:Accession: C32970
A:Molecule type: mRNA
A:Residues: 1-122, 'VF', 125-172, 'R', 174-379, 'I', 381-504 <MA1>
R:Ishida, N.; Tawaragi, Y.; Inuzuka, G.; Sugita, O.; Kubota, I.; Nakazato, H.; Noguchi,
Biochem. Biophys. Res. Commun. 156, 681-688, 1988
A:Title: Four species of cDNAs for cytochrome P450 isozymes immunorelated to P450c-M/F e
A:Reference number: A90151; MUID:89050091; PMID:3190674
A:Accession: A31579
A:Molecule type: mRNA
A:Residues: 1-122, 'VF', 125-172, 'R', 174-379, 'I', 381-504 <ISH>
A:Cross-references: EMBL:M2328; NID:g203802; PID:AAA41043.1; PID:g203803
R:Jiang, Q.; Voigt, J.M.; Colby, H.D.
Biochem. Biophys. Res. Commun. 209, 1149-1156, 1995
A:Title: Molecular cloning and sequencing of a guinea pig cytochrome P450 (CYP2D16): h
A:Reference number: J04153; MUID:95251703; PMID:7733969
A:Accession: J04158
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-504 <JIA>
R:Ohishi, N.; Imaoka, S.; Suzuki, T.; Funae, Y.
Biochim. Biophys. Acta 1158, 227-236, 1993
A:Title: Characterization of two P-450 isozymes placed in the rat CYP2D subfamily.
A:Reference number: S39761; MUID:94072607; PMID:8251521
A:Accession: S39761
A:Molecule type: protein
A:Residues: 1-9, 'X', 11-13 <OH>
C:Genetics:
A:Gene: CYP2D1
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F:1-504/Product: cytochrome P450 2D1 #status experimental <MAT>
F:4-504/Product: cytochrome P450 2D1v #status experimental <MAT2>
F:9-25/Domain: transmembrane #status predicted <TM1>
F:305-468/Domain: cytochrome P450 homology <CYP>
F:310-326/Domain: transmembrane #status predicted <TM2>
F:446/Binding site: heme iron (Cys) (axial ligand) #status predicted
Alignment Scores:
Pred. No.: 1 58e-49 Length: 504
Score: 1126.00 Matches: 342
Percent Similarity: 28.15% Conservative: 51
Best Local Similarity: 24.50% Mismatches: 100
Query Match: 6.01% Indels: 903
DB: 1 Gaps: 10
```



```
QY 5342 CATATGACATCCCGTGACATCGAAGTACAGGGCTTCGGATCCTAAGGTAGGCTGGCG 5401
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
379 ArgPheThrSerCysAspIleGluValGlnAspPheValIleProLys----- 394
QY 5402 CCCTCCTCACCCAGCTCAGCACAGCAGCCTGGTGATAGCCCGCAGCATGGCTACTGCCAG 5461
Db      ----- 394
QY 5462 GTGGGCCCACTTAGGAACCTGGCCACCTAGTCTCAATGCCACCACACTGACTGTCCC 5521
Db      ----- 394
QY 5522 CACTTGGTGGGGGTCCAGAGTATAGGCAGGGCTGGCTGTCCATCCAGAGCCCCCGTC 5581
Db      ----- 394
QY 5582 TAGTGGGAGACAAACCAGGACCTGCCAGAACTGTGGAGGACCCAGGCGCTGCAGGGAGA 5641
Db      ----- 394
QY 5642 GGGGCGAGTGTGGTGCTCTGAGAGGTGTGACTGGCGCCCTGTGTGGGTGCGAGAGGG 5701
Db      ----- 394
QY 5702 TACTGTGGAGCTTCTCGGGCGCAGGACTAGTTCACAGAGTCCAGCTGTGTGCCAGGCAGT 5761
Db      ----- 394
QY 5762 GTGTGTCCCGCTGTGTTTGGTGGCAGGGTCCCAGCATCCTAGAGTCCAGTCCCACCTC 5821
Db      ----- 394
QY 5822 TCACCCTGCATCTCTGCCCAGGGAACGACACTCATCACCACTGTCTCGGTGTGAA 5881
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
395 -----Gly-ThrThrLeuIleIleAsnLeuSerSerValLeuIly 407
QY 5882 GGATGAGGCGCTGGGAGAACCTTCGCTTCCACCCGAAACTTCTCGATGCCCA 5941
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
407 sAspGluThrValTrpGluLysProHisArgPheHisProGluHisPheLeuAspAlaG 427
QY 5942 GGGCCACTTTGTGAAGCCGAGGCTCTCTGCTTCTCAGCAGGTGCTGTGGGGAGCC 6001
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
427 nGlyAsnPheValLysHisGluAlaPheMetProPhe--Ser----- 440
QY 6002 CGGCTCCCTGTCCTCCGCTGAGTCTTGCGAGGTCTTGCGAGGGGTATCACCCAGGAGCCAGGCTCACT 6061
Db      ----- 440
QY 6062 GACGCCCTCCCTCCCCACAGGCGCGGTGCATGCTCGGGAGCCCTGGCGCCGCTAG 6121
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
441 -----AlaGlyArgArgAlaCysLeuGlyGluProLeuAlaArgMet 454
QY 6122 GAGCTCTTCTTCTTCTACCTCCCTGCTGCAGCACTTCAGTCTTCTCGGTGGCCCGCGGA 6181
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
455 GluLeuPheLeuPheThrCysLeuLeuGlnArgPheSerPheSerValProValGly 474
QY 6182 CAGCCCGGCGCCAGCACTCTCGTGTGCTGAGCTTCTGCTGACCCCATCCCCCTACGAG 6241
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
475 GlnProArgProSerThrHisGlyPheAlaPheProValAlaProLeuProTyrGln 494
QY 6242 CTTTGTGTGTCGCCCGCTAGAATGGGTA 6271
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
495 LeuCysAlaValValArgGluGlnGlyLeu 504
```

Search completed: February 25, 2004, 03:04:07
Job time : 271.885 secs

